Vol. 56, No. 6 November, 1949

Psychological Review

EDITED BY
CARROLL C. PRATT
PRINCETON UNIVERSITY

CONTENTS

Knight Dunlap: 1875–1949: KATE GORDON MOORE 309

Statistical Behavioristics and Sequences of Responses: GEORGE A. MILLER AND FREDERICK C. FRICK 311

The Problem of Quantification in Psychological Science: DANIEL BROWER 325

A Method of Graded Dichotomies for the Scaling of Judgments: FRED ATTNEAVE 334

An Application of ${}_sE_R$ Quantification Procedure: DAVID ZEAMAN 341

The Applicability of Motivational Criteria to Emotions: R. H. WATERS AND D. F. BLACKWOOD 351

The Nature and Functioning of Wants: EDWARD C. TOLMAN 357

PUBLISHED BI-MONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.
$5.50 volume $1.00 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, Section 538, P. L. and R., authorized Jan. 8, 1948

MEMBERSHIP RULES

IN THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

There are three classes of membership in the American Psychological Association: Associate, Fellow, and Life Member.

A. The largest class of membership is *Associate.* In order to qualify as an Associate an applicant must meet one of three sets of requirements:

1. He must have a doctor's degree based in part upon a psychological dissertation and conferred by a graduate school of recognized standing; *or*

2. He must have completed two years of graduate work in psychology at a recognized graduate school and be devoting full time to work or graduate study that is primarily psychological in character; *or*

3. He must have completed one year of graduate study plus one year of professional work in psychology and be devoting full time to work or graduate study that is primarily psychological in character.

Scientists, educators, or other distinguished persons who are known to and recommended by the Board of Directors may also be elected as Associates of the American Psychological Association.

Applicants must have their applications complete by August 1. New Associates are elected in the fall and their membership will be dated as of the next year. Journals due members will begin with the January issues. Annual dues for Associates are now $12.50.

B. Properly qualified Associate members may, upon nomination by one of the Divisions and election by the Council of Representatives, become *Fellows* of the American Psychological Association. Fellows must previously have been Associates. They must have a doctor's degree and at least five years of acceptable professional experience beyond that degree. They must be primarily engaged in the advancement of psychology as a science and a profession. Annual dues for Fellows are now $17.50. In the American Psychological Association, no one is made a Fellow except at his own request.

C. *Life Membership* is open to members who have reached the age of 65 and who have been members for twenty years. They are exempt from dues, and receive the *American Psychologist* and the *Directory*.

Psychological Review

EDITED BY

CARROLL C. PRATT, PRINCETON UNIVERSITY

VOLUME 56, 1949

PUBLISHED BI-MONTHLY

BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

PRINCE AND LEMON STS., LANCASTER, PA.

AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, Section 538, P. L. and R., authorized Jan. 8, 1948

LANCASTER PRESS, INC., LANCASTER, PA.

CONTENTS OF VOLUME 56

ARNHEIM, R., **The Gestalt Theory of Expression** 156
ATTNEAVE, F., **A Method of Graded Dichotomies for the Scaling of Judgments** 334
BIRCH, H. G., AND BITTERMAN, M. E., **Reinforcement and Learning: The Process of Sensory Integration** 292
BLUM, R. A., AND BLUM, J. S., **Factual Issues in the "Continuity" Controversy** 33
BROWER, D., **The Problem of Quantification in Psychological Science** 325
BURKE, C. J., **A Theory Relating Momentary Effective Reaction Potential to Response Latency** 208
CARTER, L. F., AND SCHOOLER, K., **Value, Need, and Other Factors in Perception** 200
COHEN, J., **Color Vision and Factor Analysis** 224
ELLSON, D. G., **The Application of Operational Analysis to Human Motor Behavior** 9
GATES, A. I., **Edward L. Thorndike: 1874–1949** 241
GOODENOUGH, F. L., **The Appraisal of Child Personality** 123
HARLOW, H. F., **The Formation of Learning Sets** 51
HARRISON, J. M., **Hull's Derivation of Stimulus Asynchronism: A Correction** 252
HASTORF, A. H., AND KNUTSON, A. L., **Motivation, Perception and Attitude Change** 88
HULL, C. L., **Stimulus Intensity Dynamism (V) and Stimulus Generalization** 67
JONES, F. N., **Some Psychological Implications of Cortical Suppressor Areas** 95
KELLOGG, W. N., **'Superstitious' Behavior in Animals** 172
MCGINNIES, E., **Emotionality and Perceptual Defense** 244
MASLOW, A. H., **The Expressive Component of Behavior** 261
MAY, M. A., **An Interpretation of Pseudo-Conditioning** 177
MILLER, G. A., AND FRICK, F. C., **Statistical Behavioristics and Sequences of Responses** 311
MOORE, K. G., **Knight Dunlap: 1875–1949** 309
MORLAN, G. K., **A Note on the Frustration-Aggression Theories of Dollard and His Associates** 1
OSGOOD, C. E., **The Similarity Paradox in Human Learning: A Resolution** 132
SHEFFIELD, F. D., **Hilgard's Critique of Guthrie** 284
STANLEY, W. C., AND JAYNES, J., **The Function of the Frontal Cortex** 18
TAYLOR, J. C., **Behavior Oscillation and the Growth of Preference** 77
THORNDIKE, E. L., **The Psychology of Invention in a Very Simple Case** 192
TOLMAN, E. C., **The Nature and Functioning of Wants** 357
———, **There is More than One Kind of Learning** 144
WATERS, R. H., AND BLACKWOOD, D. F., **The Applicability of Motivational Criteria to Emotions** 351
WEBER, C. O., **Homeostasis and Servo-Mechanisms for What?** 234
WOLFLE, H. M., **A Fundamental Principle of Personality Measurement** 273
WOLPE, J., **An Interpretation of the Effects of Combinations of Stimuli (Patterns) Based on Current Neurophysiology** 277
YOUNG, P. T., **Food-Seeking Drive, Affective Process, and Learning** 98
———, **Emotion as Disorganized Response—A Reply to Professor Leeper** 184
ZEAMAN, D., **An Application of $_sE_R$ Quantification Procedure** 341

Knight Dunlap

VOL. 56, NO. 6 NOVEMBER, 1949

THE PSYCHOLOGICAL REVIEW

KNIGHT DUNLAP

1875–1949

Somewhere in the writings of Taine there is a passage which emphasizes the importance of cultivating the "taste and habit of work." This striking phrase might be used as a text in describing the life of Knight Dunlap. He had the taste and habit of work. He had also the rich intellectual endowment which made his work yield important results in the fields of psychological teaching and research.

Dunlap was a native son of California, born in Diamond Spring, and it was a satisfaction to him to spend the later years of his academic service in his native state. During his training in the University of California he realized that he wanted to become an experimental psychologist. Psychology remained his dominant interest, and was pursued with a devotion and an integrity that have been the inspiration of generations of students. As we all know, Dunlap was something of an iconoclast; great names did not overawe him. But the name of the man who revealed to him the meaning and the possibilities of experimental psychology was one which he always mentioned with a respect not short of reverence. That was the name of George Malcolm Stratton.

After completing his doctorate at Harvard, Dunlap returned to Berkeley as instructor in psychology. From Berkeley he was called to The Johns Hopkins University where he served from 1906 to 1936, and held the title Professor of Experimental Psychology during the last 20 years. During 1917–1918 he served with the Army in the Air Medical Service. In 1936 he joined the faculty of the University of California at Los Angeles as Professor and Chairman of the Department of Psychology, and remained there until his retirement in 1946.

In addition to his teaching and research Dunlap influenced the development of psychology by way of his editorial activities. He founded the *Journal of Psychobiology* in 1917, and when that journal merged with the *Journal of Animal Behavior* to form the *Journal of Comparative Psychology*, the editorial responsibility was largely his. Other journals and monograph series likewise benefited by his help.

Dunlap's publications were numerous. His first article "The effect of imperceptible shadows upon the judgment of distance" appeared in this JOURNAL in 1900. His first book was *A System of Psychology*, published in 1912, and the last, in 1946, was *Religion, Its Functions in Human Life*. He was responsible for many experimental studies, several well known books, and the invention and design of apparatus.

It is not, however, the quantity of a man's output for which we remember him. Dunlap was original and stimulating to a high degree. Fortunately we have his own summary of his main

contributions, and those who are students of psychological literature will recognize the central importance of the problems with which he wrestled. He classifies his work as theoretical, experimental and instrumental, and under the caption of theory we have the following seven topics.

1) The attack on introspection. 2) The insistence on response or reaction as the basis of mental processes, including thought processes. 3) The attack on images and the insistence on objects of perception and of thought, in general, as real objects, instead of "psychic" objects, "states of consciousness," "sensations," and "images." 4) The emphasis on the periphery, in general, instead of on the brain, as the fundamental determiner of psychological qualities. 5) The elimination of "instincts" and the emphasis on desires. 6) The view of consciousness as an abstract reference to an undeniable fact of response. 7) The reformation of our views on heredity.

In elaborating these positions he made many forceful and subtle comments. Whether one agrees or not with his theories there is no question but that he stimulated the discussion and advanced the understanding of vital psychological problems. In later years he gave much time to social, religious and clinical studies. In discussing child care he once said that "keeping the baby happily stimulated during its waking periods, and preventing crying, while not 'spoiling' the child, is a difficult task, too difficult, perhaps, for the intelligence of many parents. Spoiling a child, however, is a minor evil; neglecting it is a major one."

Among his many memorable sentences the following could be taken as an important part of his creed. "As a psychologist, I have the psychologist's prejudice, that ideals, intellectual analysis, and education are the fundamental forces of progress, and that laws, conventions, and customs serve to consolidate and make secure the gains achieved through these forces."

Knight Dunlap was a man most generous of his time and thought, keen in analysis and criticism, and fertile in psychological invention. For the ten talents lent to him he made a rich return.

KATE GORDON MOORE

STATISTICAL BEHAVIORISTICS AND SEQUENCES OF RESPONSES

BY GEORGE A. MILLER AND FREDERICK C. FRICK

Harvard University

This paper is concerned with procedures that can be used to analyze serial dependencies in chains of responses. In order to introduce these procedures we shall begin with some general opinions about experimentation in modern psychology. These opinions comprise a point of view which we propose to call *statistical behavioristics*. This point of view represents a methodological bias that we believe is accepted by many experimental psychologists; it has not, however, received explicit formulation nor has it been fully exploited.

We have chosen to call this approach *behavioristics* because behavior seems to be the observable datum of psychology. We have modified the noun by the word *statistical* because behavior has that kind of complexity that seems to preclude, at least for the present, a completely deterministic account.

Statistical Behavioristics

Most psychological experiments involve the classification of an organism's responses and the tabulation of the number of responses in each class. There are many difficulties concealed in this summary sentence—the decisions as to which responses to record, which environmental changes to use in their classification, how large a sample is necessary, etc. In what follows it will be assumed that the psychologist has been able to limit his interest to certain portions of the behavior stream that he feels are peculiarly diagnostic, and that these portions can be recognized whenever they occur. They have the status of "responses" or "response units." The psychologist's problem is to describe the relative frequencies of occurrence of responses in each of the classes he has set up.

A psychophysical example. In order to make the discussion more concrete, consider an example from psychophysics. The subject pushes one key when a stimulus increases in magnitude, another key when the stimulus decreases in magnitude. All the subject's responses during the session are not recorded, but only the subset of responses that involve pushing a key. This subset of recorded responses, R, can be classified in two ways: (1) into subset A, the responses to key a, and subset B, the responses to key b, or (2) into subset S_1, the responses to stimulus change s_1, subset S_2, the responses to stimulus change s_2, and subset S_3, the responses to stimulus change s_3. In order to correlate responses to both key and stimulus, further subsets are formed by the logical products (those elements common to the two subsets): $A \cap S_1$, the set of responses to key a when s_1 was presented, $A \cap S_2$, the set of responses to key a when s_2 was presented, etc. The final step is to infer the differential limen from the relative frequencies of responses in each of the subsets.

The classification is shown graphically in Fig. 1. Fig. 1A represents each response by a dot, and responses included in each subset are surrounded by the subset contours. Fig. 1B gives the classification in the form of

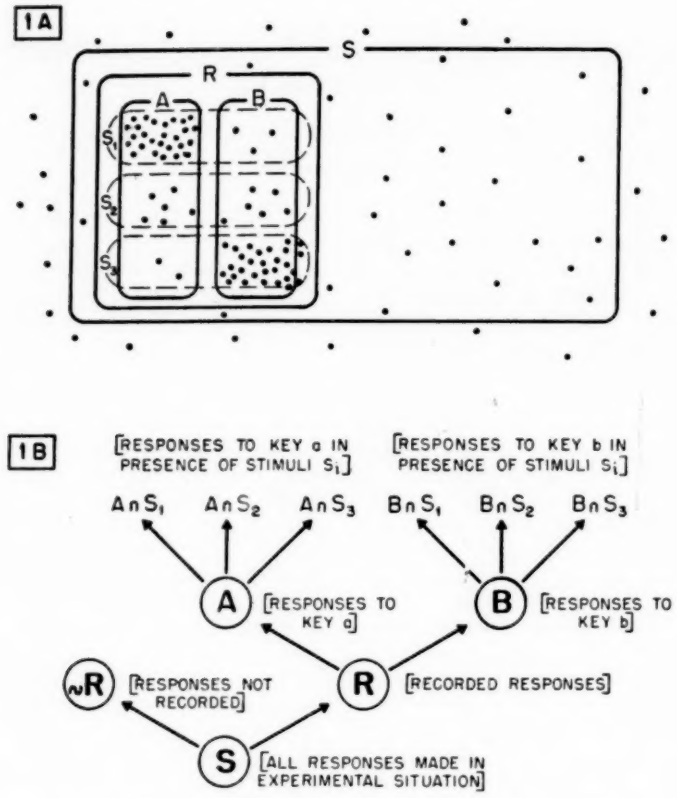

FIG. 1. Two graphical representations for the classification of responses in a hypothetical psychophysical experiment.

a tree, where the classes become more specific as we ascend the tree. The classification can also be given by the headings on a statistical table that summarizes the data (*cf.* Table I).

The classification is fundamental because it must contain, along with the definitions of the classes, all the information that is available for the analysis of the experiment. If the classification is complete, there is no further information about any element of *S*. The classification says that any element of *S* possesses certain specified properties, and to that extent the information narrows the search for the particular element to the set of all those elements of *S* that possess these properties. If the results of an experiment can be presented in a table they can be described in these general terms. The basic observations are concerned with the relative frequencies of responses in different classes.

It may be convenient to summarize the distribution of responses in the different classes by a single statistic and to describe how this statistic varies with other dimensions of the

situation. This extension complicates the classification rather than changing its fundamental nature, and usually neglects some of the information.

The course of action. It is important to notice that the psychophysical classification is unordered. Responses were classified first on the basis of the key depressed, then on the basis of the stimulus presented. Classification in the reverse order would have served the purpose equally well. Unordered classifications have been the rule in psychological research, and most of the familiar statistical procedures are designed to handle such classifications.

There are situations, however, where the order of classification is definitely specified. For example, geneticists build classifications that include the relation "parent of," and successive classes cannot be interchanged.

An ordering relation that is of particular interest to the psychologist is "precedes in time." This relation gives sequential information that has been largely neglected in the classifications of responses. In the psychophysical experiment just described, for example, the successive judgments were embedded in the time continuum, but no record of the order was kept and the serial information is not available in the classification shown in Fig. 1.

In order to include the sequential information, the first step is to decide what responses to record. In Fig. 1 the recorded responses are defined according to key and stimulus intensity into six subsets—refer to these now as responses A, B, C, D, E, and F. The next step is to distinguish the set of responses of set A that were immediately followed by responses of set A; the set of responses of set A that were followed by responses of set B; the set of responses of set A that were followed by responses of set C; and so on until all possible digrams are listed. In this example there will be 36 digrams.

By distinguishing the various possible pairs of responses another step is added to the classification of Fig. 1. As many further steps can be added as necessary—just how far we need to proceed will be discussed later. The set of recorded responses is partitioned into A, B, C .... Each of these classes is partitioned into AA, AB, AC,... BA, BB, BC,...CA, CB, CC, etc. Then each of these is further partitioned into AAA, AAB, AAC,... ABA, ABB, ABC,... ACA, ACB, ACC, etc. Sequences of four, five or n responses in a row can be tabulated to preserve the sequential information.

Ignoring the sequential information is equivalent to assuming that successive responses are independent. If this assumption is correct there is little information to be gained from a sequential tabulation. According to the assumption of independence, if the relative frequency of response A is $p(A) = 0.5$, and of response B is $p(B) = 0.5$, we could assume that the relative frequencies of AA, of AB, of BA, of BB are all 0.25, and of triads are 0.125, etc. If an actual tabulation of sequences shows significant departures from these values, then the assumption of independence is not justified, and the frequency distribution for sequences of n responses cannot be derived from the distribution for individual responses.

It is easy to see that a sequential classification can be exhaustive. To take a specific example, suppose the experiment defines two responses, 0 and 1, and that the sample consists of sequences of ten such responses from each subject. The possible number of different sequences of ten responses

with two alternatives is 2^{10}, or 1024. If the sequential classification is continued out to sequences of ten responses, the classification will include all the 1024 *courses of action* that are possible in the experimental situation.[1]

In summary: Psychologists define classes of responses, determine their relative frequencies of occurrence and attempt to discover how the frequency distributions change as the experimental variables are manipulated. Within this broad area of statistical behavioristics lie problems that involve sequences of responses, or courses of action. The problem is to predict the frequencies of these courses of action, just as individual response frequencies are predicted in the nonsequential case.

Approximate Representation of the Course of Action

The psychologist's experiments usually generate a sequence of symbols: right and wrong, conditioned and unconditioned, left and right, slow and fast, adient and abient, etc. The analysis must begin with this system of symbols that the experimenter writes in his records. The record might consist of *R*'s for correct responses in a learning experiment and *W*'s for errors, and the analyzer would be faced with something like this:

RWWWWRWRWR-
WWWRWRWWWR, etc.

In the first twenty responses there are 0.35 correct responses and 0.65 errors.

Now consider the mathematical models that might be taken from the theory of probability and used to summarize these observed frequencies. In the simplest model each symbol in the sequence leaves the probabilities unaltered. The occurrence of the symbol returns the system to the same state, and the distribution of probabilities is the same after the symbol occurs as it was before. This formal model is presented graphically in Fig. 2. Here the arrows represent the occurrence of a symbol, the dot represents the single state of the system, and both arrows return the system to the original state.

FIG. 2. Diagrammatic representation of a sequence of independent events. An arrow is traversed when the event occurs, and both events return the process to the same original state.

Dependent probabilities. The formal model depicted in Fig. 2 considers successive symbols as independent events. This model would describe the results obtained by drawing marbles (and replacing them) from an urn that contained the appropriate proportions of *R* and *W* marbles.

So simple a model may not be sufficient. It may be necessary to consider sequential dependencies, which is done by counting the occurrences of pairs of items. For the record given above such a count shows seven occurrences of *WW*, six occurrences of *RW*, six occurrences of *WR*, and no occurrences of *RR*. The results can be presented in the form shown in Table I.

When confronted with four-fold tables of this sort, psychologists naturally turn to a contingency measure of the degree of association between the arguments. Since these

[1] A sequence given in 0s and 1s can be treated as a number in the binary system and translated to the corresponding number in the more familiar decimal system. For example, 0001011101 is the equivalent of 93. Thus the decimal numbers from 0 to 1023, when jotted on the record, preserve the complete sequential information in the example discussed.

TABLE I

SHOWING THE DISTRIBUTION OF PAIRS OF RESPONSES

		First response		Total
		Right	Wrong	
Second response	Right	$p(RR)=0$	$p(WR)=6/19$	$p(R)=6/19$
	Wrong	$p(RW)=6/19$	$p(WW)=7/19$	$p(W)=13/19$
Total		$p(R)=6/19$	$p(W)=13/19$	1.0

techniques are familiar, the following discussion will be restricted to the novel, but related, methods of analysis that have been developed as a part of the mathematical theory of communication (4).

Note that the relative frequencies of R and W are not the same when R has just occurred as they are when W has just occurred. If this difference is to be reflected in the mathematical model, the model must have two distributions of probabilities—one distribution that applies following the symbol R, and another distribution that applies following the symbol W.

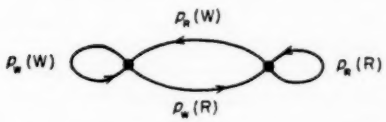

FIG. 3. Diagrammatic representation of the second-order approximation of a sequence of dependent events. The event R always leaves the process in one state, while the event W leaves it in the other.

Or, phrased differently, the system must have two states, one state that follows R and another state that follows W. The occurrence of a symbol does not always return the system to the same state, and thus the mathematical model admits dependent probabilities. This formulation is shown graphically in Fig. 3.

As in Fig. 2 the dots in Fig. 3 represent the states of the system, and a symbol is generated whenever an arrow is traversed. There are several probabilities to be stated for the model in Fig. 3. There are the probabilities that the organism will get into one state or the other, $p(W)$ and $p(R)$. There are the probabilities of the several pairs of responses, $p(WW)$, $p(WR)$, $p(RW)$ and $p(RR)$. And there are the transitional probabilities, $p_W(W)$, $p_W(R)$, $p_R(W)$ and $p_R(R)$, that when the organism is in a given state (indicated by the subscript) the next response will leave the organism in the same state or change it to a new state (indicated by the argument). All these probabilities are related in a straightforward way:

$$p(W)+p(R)=p(WW)+p(WR)+p(RW)+p(RR)=1,$$
$$p_W(W)+p_W(R)=p_R(W)+p_R(R)=1,$$
$$p(WW)=p(W)\cdot p_W(W).$$

The mathematical systems represented in these diagrams are called finite, discrete Markoff processes, a special class of stochastic processes. A finite number of possible states of the process exist, Q_1, Q_2, Q_3,...Q_r. A different distribution of probabilities holds for each of the several states, and which distribution applies depends upon the preceding events that got the system into its present state. Thus the states represent the residual

influence of the preceding events. A sequence of infinite length can be generated by a process with a finite number of states, for the system can return to any state over and over again.

A sequence of events can have interdependencies extending over several successive events. If such a series is represented by a Markoff process of the type shown in Fig. 2, it is a first-order approximation. If the probabilities of pairs of events are considered, as in Fig. 3, it is a second-order approximation. When the analysis is extended to sequences of s events, it is the s-th-order approximation. Most psychological experiments have not been analyzed beyond a first-order approximation, where successive responses are assumed to be independent.

In general, if the classification distinguishes a mutually exclusive events, the s-th-order approximation is made up of a^s transitional probabilities $p_{ij\ldots}(k)$ (the probability that, given the sequence $ij\ldots$, the event k will follow), and a^{s-1} states corresponding to the various sequences of s-1 events that might precede the s-th event.

To summarize:

a^s	The number of *classes* necessary for a complete tabulation of a alternatives considered in sequences of s.
a^{s-1}	The number of *states* of the Markoff process, the number of tabulated antecedents of the s-th event in the sequence of s.
$a^{s-1}(a\text{-}1)$	The number of *degrees of freedom* in the tabulation. There are a^{s-1} different states, and each state has associated with it a-1 degrees of freedom.

These mathematical models generate infinite sequences of events. This would seem to rule out the Markoff process as a model for the study of learning, for during the course of learning the relative frequencies change progessively.[3] This is not a new problem to the psychologist, for whenever statistical models based upon hypothetically infinite populations are applied to behavior he is faced with the sampling problem. The samples must be large enough to allow the relative frequencies to approach their limits, but short enough to insure that the process has not changed significantly from the beginning to the end. It should be noted, however, that the sampling difficulty is acute with the type of statistical model discussed here. For example, in a third-order approximation with three alternative responses the "behavior space" is granulated into 27 classes. The proportion of responses in each class must have time to approach the parametric value. If the sample consists of 20 responses (18 sequences of three responses), it is clearly impossible to get an adequate sample for each of 27 classes.

If the behavior system changes appreciably on each successive response, elaborate statistical description is limited. When the behavior has settled into a persistent pattern, large samples can be obtained before the system changes, and extensive analysis is possible.

From the consideration of dependent probabilities it is clear that the extensive form of classification, discussed in the first section of this paper,

[3] Dr. Frederick Mosteller has suggested informally that a broader interpretation is possible. Learning might be regarded profitably in terms of a higher-order process of the Markoff type that has Markoff processes as its component states.

can be considerably abbreviated and still be descriptively adequate. To review briefly: all the possible courses of action in a given experimental situation can be represented by an extensive classification based upon sequential information. The problem is to know which courses of action are most frequent under a given set of circumstances. If the successive responses can be regarded as generated by a Markoff process, the probability for any complete course of action can be summarized in a relatively small number of transitional probabilities.

AN INDEX OF BEHAVIORAL STEREOTYPY

The next question concerns the experimenter's accuracy in predicting the successive responses of his subject. Predictability depends upon (1) the range of alternative responses the subject can make and (2) the relative frequencies with which these alternatives occur. A quantitative index of predictability, or, as we shall call it here, an *index of behavioral stereotypy*,[3] must be based upon these parameters. Such a statistic can be derived for the mathematical models discussed above (4).

The number of alternatives. Consider first the case where all classes of events are equally probable. If there are only a few alternative classes a prediction is more certain than if there are many alternatives. "Anything can happen" is a common expression of uncertainty. In other words, the uncertainty U is a monotonic increasing function of the number of alternatives. Some freedom is possible in the choice of this function, but a logarithmic function is the most convenient.

This convenience can be demonstrated with a simple numerical example. Suppose a choice is made from among 10 equally probable alternatives, and that, having made this first choice, a second choice must also be made from among 10 equally probable alternatives. Two successive choices from 10 alternatives make possible 100 alternative pairs of choices—when the alternatives are equally probable and successive choices are independent, two choices from 10 alternatives are equivalent to one choice from 100 alternatives. If the uncertainty in these two situations is represented by a linear function of the number of alternatives, the first choice from 10 alternatives represents 10 units of uncertainty, and the second choice adds 90 more. It is not obvious why the second event should be assigned nine times the uncertainty of the first, and it would be more satisfactory if a choice from 10 alternatives always represented the same amount of uncertainty. In other words, we would like to have an additive measure of uncertainty. This is realized if uncertainty is a logarithmic function of the number of alternatives. Then the uncertainty of the first event is log 10 and the uncertainty of the second event is log 10; the total uncertainty for two successive events is 2 log 10, which is equal to log 100, the uncertainty of one choice from 100 alternatives.

In general, if there are a equally probable ways an event can happen, and if it occurs s successive times, U is constant so long as a^s is constant. It is desirable to have the uncertainty of a^s possibilities, $U(a^s)$, equal to s times the uncertainty of a alternatives.

[3] Not the best possible term, perhaps. We use it to refer to the "orderliness" of behavioral sequences, without specifying whether the orderliness is due to learning, fixation or lack of understanding. Just how this usage is related to that current among social psychologists is a matter for some debate.

That is to say, it is convenient if

$$U(a^s) = s \cdot U(a).$$

Consider two equally uncertain situations, where $a^s = b^t$. Take the logarithm of both sides of this equation,

$$s \log a = t \log b.$$

For the two equally uncertain situations,

$$s \cdot U(a) = t \cdot U(b).$$

Substituting for t/s and solving for $U(a)$:

$$U(a) = \frac{U(b)}{\log b} \log a$$
$$= K \log a = -K \log (1/a).$$

The uncertainty involved in a choice among a equally probable alternatives is proportional to the logarithm of a.

Disproportionate probabilities. The preceding discussion is restricted to alternatives of equal probability. This formulation can be extended to include alternatives of different probabilities. A numerical example will illustrate the extension. Suppose the choice is from three alternatives that have probabilities of 0.2, 0.3 and 0.5. This can be transformed into an equivalent case with equal probabilities, for it is possible to imagine that this is the first of two successive choices, and since the imaginary second choice is at our disposal it can be used to make all the pairs of choices equally probable. After choosing $p_1 = 0.2$ there would then be a choice between two equally likely alternatives; after $p_2 = 0.3$ a choice among three alternatives, and after $p_3 = 0.5$ a choice among five alternatives. The compound probabilities for the pairs of choices are then 0.1, and the uncertainty of the double choices is equivalent to the uncertainty of a single choice among 10 equally probable aternatives. This situation is illustrated in Fig. 4.

Now calculate the uncertainties involved in Fig. 4. In addition to the uncertainty of the first choice, $U(\frac{2}{10}, \frac{3}{10}, \frac{5}{10})$, 0.2 of the time there is a second choice of uncertainty $U(\frac{1}{2}, \frac{1}{2})$, 0.3 of the time a second choice of $U(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, and half the time a second choice of $U(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$. Since we want an additive measure of uncertainty, we are interested in the sum of all these terms. The sum is equal to the uncertainty of a single choice among 10 equally probable alternatives. This can be written as

$$U(\tfrac{2}{10}, \tfrac{3}{10}, \tfrac{5}{10}) + 0.2\,U(\tfrac{1}{2}, \tfrac{1}{2}) + 0.3\,U(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}) + 0.5\,U(\tfrac{1}{5}, \tfrac{1}{5}, \tfrac{1}{5}, \tfrac{1}{5}, \tfrac{1}{5}) = U(\tfrac{1}{10}, \tfrac{1}{10}, \ldots \tfrac{1}{10})$$

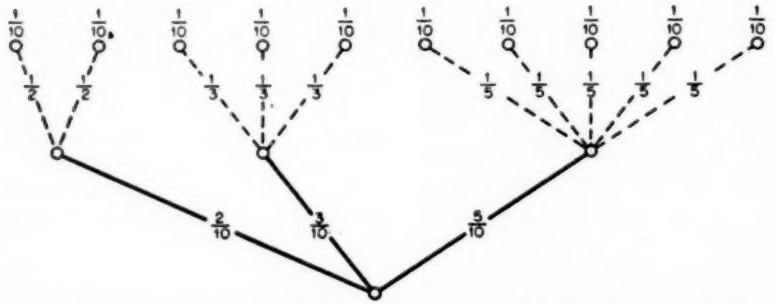

FIG. 4. Showing how a choice among unequally probable events is transformed by a second choice into a choice among equally probable pairs of events.

Four of the five terms in this equation involve equally probable alternatives with determinable uncertainty. So

$$U(\tfrac{2}{10},\tfrac{3}{10},\tfrac{5}{10}) = K \log 10 - 0.2K \log 2 - 0.3K \log 3 - 0.5K \log 5.$$

With logarithms to the base 10 the answer is

$$U = K(1 - 0.06 - 0.14 - 0.35) = 0.45K$$

Note that if the three alternatives had been equally likely, U would have been K log 3, or $0.48K$. *Any departure from equal probabilities reduces the uncertainty.*

The procedure used in the example can be generalized to give a solution for any set of disproportionate alternatives. In a general notation $p_i = n_i/\Sigma n_i$, where p_i is the probability of an event occurring in manner i, and n_i is the number of occurrences of members of class i. Then

$$U\left(\frac{n_1}{\Sigma n_i}, \frac{n_2}{\Sigma n_i} \cdots \frac{n_a}{\Sigma n_i}\right) + K\Sigma \frac{n_i}{\Sigma n_i} \log n_i = K \log \Sigma n_i,$$

where the second term on the left and the term on the right side of the equation represent the uncertainties for the equally probable alternatives that are used to obtain the solution. Now solve the equation for U:

$$U = K(\log \Sigma n_i - \Sigma p_i \log n_i).$$

Multiply the constant log Σn_i by $\Sigma p_i = 1$, to obtain

$$U = K(\Sigma p_i \log \Sigma n_i - \Sigma p_i \log n_i)$$
$$= -K\Sigma p_i \log \frac{n_i}{\Sigma n_i}$$
$$= -K\Sigma p_i \log p_i$$

This derivation is given for the field of rational numbers, but if the p_i are in the field of irrationals, they may be approximated by rationals. This approximation requires the assumption that probability is a continuous function.

An examination of the formula for U shows that U is the mean value, or mathematical expectation, of the function log p_i. In other words, when the probabilities of the separate alternatives are not equal, we use the expected value of log p_i.

The derivation gives a general expression for uncertainty which is equivalent to the physicist's equation for entropy. When the probabilities for the various alternatives are known, this formula can be used to compute the uncertainty with which any particular, individual response can be predicted in the absence of any knowledge about the preceding or following responses.

The value of K in this equation amounts to a choice of a unit for the measure of uncertainty. It is convenient to choose K in such a way that it equals 1 when the base of the logarithm is taken as 2. If the base 2 is used the resulting units are binary digits, or more briefly *bits*. The maximum uncertainty in predicting one of two alternatives is one bit of uncertainty. Making s such decisions in a row gives s bits of uncertainty, since the total number of possible alternatives is 2^s, and $\log_2 2^s = s$. Of course logarithms to the base 10 can be used, and the resulting units are decimal digits. A decimal digit is 3.32 bits. To express the equation in bits when logarithms to the base 10 are used, K must have the value 3.32. In the following discussion we will use the bit as the unit of uncertainty.

Dependent probabilities. It has been assumed that a behavior sequence can often be represented by a Markoff process with a definite number of states. If the process is in any given state there are several possible

events that can follow. Consequently, there is a degree of uncertainty associated with each state. In one state the next event may be very certain, while in another state the next event may be relatively uncertain. To represent the uncertainty for any particular state, write $-\Sigma_j p_1(j) \log_2 p_1(j)$, where $p_1(j)$ is the probability that state 1 will be followed by event j, and the summation is for all the possible j that could occur. The uncertainty for the whole process must be the weighted average of the uncertainties for each of the separate states. The contribution of each state to the average is weighted according to the probability of occurrence of the state, P_i, and so represents the mean value or mathematical expectation of U for the several states. Thus the average uncertainty per response, when the s-th approximation is used, is

$$U_s = -\sum_{i,j} P_i\, p_i(j) \log_2 p_i(j).$$

This equation represents the average uncertainty in the prediction of the next event when the preceding s-1 events are known.

This equation is simplified by taking advantage of the fact that $P_i p_i(j)$ represents the probability that the sequence leading to state i and also event j will occur together. This probability is symbolized by $p(Q_i,R_j)$, the probability that state (sequence) Q_i followed by event R_j will occur. Similarly, the probability that, given sequence Q_i,R_j will follow is indicated by $p_{Q_i}(R_j)$. The equation is written as follows:

$$U_s = -\Sigma p(Q_i,R_j) \log_2 p_{Q_i}(R_j).$$

A series of approximations to U is obtained by considering the statistical structure of sequences of 1, 2, 3...s events. Shannon (4) has shown that U_s is the uncertainty of the sth order approximation, and as s increases without limit, U_s approaches U. Further, U_s is a monotonic decreasing function of s, and the uncertainty of the next event when the preceding s-1 are known is never increased by a knowledge of the statistical structure of events preceding that. Analysis of sequences of events longer than s cannot make the prediction less certain, but may make it more certain.

When the mathematical model is used to summarize behavioral data, U_s represents the uncertainty with which a response can be predicted when only the preceding s-1 responses are known. It will be called the *uncertainty per response.*

To compare results from different experiments it is often convenient to use the *relative* uncertainty per response. For any situation the maximum value of U is computed by assuming all responses to be independent and equally frequent. If there are two alternative responses, the maximum uncertainty will be $-2(0.5 \log_2 0.5)$, or 1.0. With four possible responses the maximum uncertainty is $-4(0.25 \log_2 0.25)$, or 2.0. In general, $U_{max} = -\log_2(1/a) = \log_2 a$, where a is the number of possible alternatives. The relative uncertainty, therefore, is given by the ratio of U_s to U_{max}, and this ratio may vary between 0 and 1.

We are now in a position to define an *index of behavioral stereotypy*:

$$C_s = 1 - (\text{relative } U_s)$$

A value of C_s of 0.50, for example, is equivalent to the statement that half of the responses are determined and that half are maximally uncertain.

AN EXPERIMENTAL ILLUSTRATION: MULTIPLE-CHOICE BEHAVIOR

The discussion in the preceding sections can be crystallized in an experi-

mental example. In 1916 G. V. Hamilton (1) reported a study of trial-and-error behavior in a multiple-choice situation and, fortunately, published complete data for many of his subjects.[4] Hamilton used an apparently unsolvable multiple-choice problem that required the subjects to find their way out of an enclosure from which there were four possible exits. The correct exit varied from trial to trial, so it was useless to try again the exit that had proved successful on the preceding trial. No other clue was available. This apparatus was supposed to prevent the formation of a specifically adaptive habit and to provide an extensive sample of trial-and-error behavior.

The question to ask of these data is how unpredictable or how stereotyped is trial-and-error behavior in this situation? Does a subject systematically vary his responses to produce a random sequence of exit attempts, or does he settle down to a relatively routine and stereotyped pattern that is adequate though not always rewarded?

When the data from all of Hamilton's 7-year-old girls are taken together the sample is large enough to permit reasonable analysis. As a sample of the records, successive exit attempts might run as follows, where the numbers designate the four possible exits:

1, 2, 3, 4, 3, 2, 1, 4, 3, 2, 3, 4, 2, etc.

The record was tabulated for sequences of three successive responses. The first three responses constituted an entry under the (1, 2, 3) category, responses two through four made an entry in the (2, 3, 4) category, responses three through five made an entry in the (3, 4, 3) category, etc., until the day's trial was completed, *i.e.*, until the exit had been found ten times. When four alternatives are tabulated in sequences of three responses, there are $4^3 = 64$ different categories. The data for the five girls consisted of slightly over 1000 responses and so permitted a fair sample for all 64 possible sequences. A tabulation of sequences of four, however, requires 256 possible classes, and Hamilton's sample is too small for reasonable reliability.

From the tabulation of sequences of three responses we can, of course, determine the tabulation for sequences of two or one, since $p(1, 2) = p(1, 2, 1) + p(1, 2, 2) + p(1, 2, 3) + p(1, 2, 4)$, etc. The tabulation for the longer sequence contains all the information given by tabulations of shorter sequences.

The first approximation for the behavior of the five girls shows the proportion of responses to door 1, $p(1)$, to be 0.21, $p(2)$ to be 0.31, $p(3)$ to be 0.30, and $p(4)$ to be 0.18. For individual responses the distribution is almost uniform, with only a slight tendency for the girls to favor the middle two doors. When there are four alternatives, the maximum value of U is $\log_2 4$, or 2 bits. The calculated value of U_1 for these data is 1.96. The relative uncertainty is, therefore, 1.96/2, or 0.98. The index of stereotypy for the first approximation is $C_1 = 0.02$. With no knowledge of preceding responses any prediction of the next exit to be tried is very uncertain.

When sequences of two responses are considered, however, the picture is quite different. Of the 16 possible ordered pairs of responses, six predominate: (1, 2), (2, 3), (3, 4), (4, 3), (3, 2) and (2, 1). These six response

[4] It surprised us to discover how few psychological studies have reported the actual sequential information necessary for this analysis. It seems to be the first information that is thrown away when the data are summarized for publication.

pairs together comprise 80 per cent of the tabulated pairs. They are made up of responses to adjacent exits—apparently it is unusual for a subject to walk past a door without stopping to try it, or to try the same door twice in succession. These tendencies introduce a certain degree of orderliness into the behavior, and C_2 is 0.27. With knowledge of only one preceding response it is evident that trial-and-error behavior in this situation is not unpredictable, but is stereotyped and relatively orderly.

It is interesting to compare the seven-year-old girls with the single adult rat for which Hamilton published complete data. Inspection of the tabulations of pairs of responses shows a striking similarity between girls and rat. The same six pairs of responses that the girls favored were favored by the rat, and comprised 78 per cent of its responses. For the rat C_1 was 0.06, and C_2 was 0.26. On the basis of a second-order approximation it is difficult to discover any difference between the animal's and the girls' adjustment to the situation. This in spite of the fact that it took the rat approximately twice as many responses to complete 100 successful exits as it took the average girl. The better adjustment of the girls is not revealed in the preceding analysis, and it is necessary to consider the third-order approximation before the difference becomes clear.

The sequences of three responses show that the girls tended to make a third response different from either of the two preceding responses. There was no need, they discovered, to try again a response that was unsuccessful the time before last. This tendency appears in the form of greater stereotypy in their behavior, and C_3 is 0.42. The rat, on the other hand, was not as proficient in remembering what he had done on the trial before the last, and often fell into the unproductive pattern of which 3, 2, 3, 2, 3, 2, etc., is an example. If a girl made the responses 1, 2 it was almost certain that her next response would be 3, but the rat was no more apt to go on to 3 than he was to return to 1. C_3 for the rat, therefore, was only 0.31. It is unfortunate that too few data are available for a fourth-order approximation, for the indications are that the difference between girls and rat would be even more apparent.

In Fig. 5 C_s is plotted as a function of s. It is clear that the girls respond to remotely preceding responses to a greater extent than does the rat. Thus the index of stereotypy gives a second important fact about the behavior. Not only can we say how stereotyped the response sequence is, but we can estimate the length of the chain over which preceding responses have an effect. If one agrees with Hunter (**2**) that a response influenced by the penultimate of the preceding responses indicates a symbolic process, then Hamilton's experiment shows that girls have a greater capacity for symbolic processes than do rats. This is not a startling conclusion, but it is comforting to find that the computations do not contradict the obvious.

THE STATISTICAL STRUCTURE OF BEHAVIOR

The central idea in the preceding discussion is the course of action. If interest is confined to individual responses and no attempt is made to describe and discuss sequences of responses, experimental psychology may have difficulty in meeting many problems posed in the clinical and social areas. Some such concept as the course of action seems inevitable.

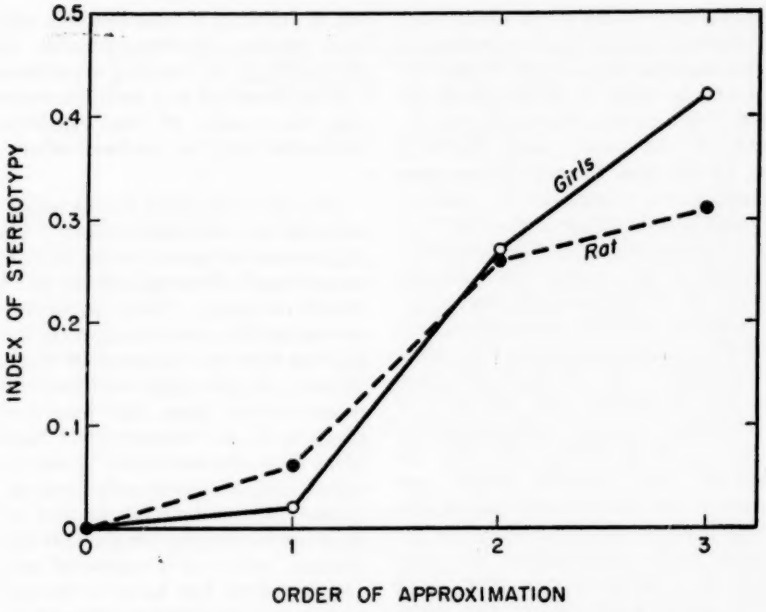

FIG. 5. The stereotypy of trial-and-error behavior. Data from Hamilton (1).

To define a course of action it is necessary to consider all the alternative courses of action that might occur. It is with this set of alternatives—this classification—that the behavior analyst must work. He is faced, not with individual responses or even with individual sequences of responses, but rather with a matrix of alternatives. His job is to decide which ones of this large set of alternatives are most frequent, then to discover the rules for manipulating these frequencies. The problem is, therefore, to analyze the statistical structure of the set of all possible courses of action. Or, more generally, to analyze the statistical structure of behavior.

This line of thinking makes it obvious that it is not necessary for a scientific psychology that one course of action be very frequent and all others very infrequent. Such stereotyped behavior seems unadaptive in most situations. The job is to describe precisely the *distribution* of frequencies among the various alternatives. C_s, the index of stereotypy, may be zero, but if it is determined carefully enough and the conditions that make it zero are thoroughly understood, we have the kind of predictability necessary for a science of behavior.

In different studies the interest may be in particular aspects of the statistical structure, or in changes in particular aspects. In learning studies, for instance, the criterion may often be set in terms of the length of the sequence, s, that can become interdependent. Learning to talk is perhaps the best example of such a behavioral change. The young child is capable of producing all the sounds involved in the language, but his job is to learn the useful sequences. The

predictability of any given sound may not change appreciably, but the length of the response sequence that must be analyzed in order to give a close approximation to the behavior is greatly extended. In other cases, learning may be directed toward greater predictability—a change in C—without any significant change in the length of the sequence that must be analyzed. Learning to hit a target, for example, is a matter of increasing the stereotypy of behavior without changing markedly the response sequence that must be analyzed. In still other situations the goal is to increase the rate of production without modifying either C or s. Typewriting provides an example. Any literate person can generate the printed symbols of language if given sufficient time, but the expert typist accomplishes the same result in a much shorter time. The final products are nearly the same, and the measure of learning is the rate of production of behavior having a specified structure. Fourthly, there is the adjustment to an altered environment. Such adjustment often involves the breaking of established patterns of behavior and the adoption of new patterns that more nearly maximize the gratification of needs. Such learning needs to be described in terms of the specific shifts in the frequencies of the several courses of action. The values of C, s and the rate of responding may remain constant although the particular dependencies are altered.

We have too long classified types of learning according to the particular experimental procedure involved in measuring the learning. Whether or not all learning is reducible to a single basic process, it seems possible that the results of all learning experiments can be described in a uniform manner and the results of one experiment compared with the results of others.

The ideas discussed in this paper do not originate with the authors. Their origins can be traced to the literature of communication engineering and the theory of games. Most important in the realm of communication engineering has been the influence of Norbert Wiener (**5**), although we have drawn more heavily upon the formulation given by C. E. Shannon (**4**). Similar ideas have also been used in von Neumann and Morgenstern's analysis of games (**3**), and their concept of a strategy is directly comparable to the present notion of a course of action. Our attempt has been to formulate certain psychological problems in such a way as to take advantage of the techniques developed by these mathematicians.

REFERENCES

1. HAMILTON, G. V. A study of perseverance reactions in primates and rodents. *Behavior Monogr.*, 1916, **3**, No. 2.
2. HUNTER, W. S. The behavior of raccoons in a double alternation temporal maze. *J. genet. Psychol.*, 1928, **35**, 374–378.
3. NEUMANN, J. v., & MORGENSTERN, O. *Theory of games and economic behavior.* Princeton: Princeton Univ. Press, 1947.
4. SHANNON, C. E. A mathematical theory of communication. *Bell System Technical Journal*, 1948, **27**, 379–423, 623–656.
5. WIENER, N. *Cybernetics.* New York: John Wiley, 1948.

[MS. received May 13, 1949]

THE PROBLEM OF QUANTIFICATION IN PSYCHOLOGICAL SCIENCE [1]

BY DANIEL BROWER

New York University

Under the influence of the behavioristic movement in American psychology, the quantitative approach has become a *sine qua non* for psychological theory, research, and practice. If we assume that the task of science is to establish hypotheses which enable us to describe, explain, predict, and control certain classes of phenomena, then the question before us is: To what extent does quantification facilitate these aims of psychological science and to what extent does quantification impede the attainment of these goals?

Quantification refers to the process of expressing our observations of phenomena in terms of numbers or other discrete figures. It would seem helpful to clarify at the outset what we mean by quantification in this sense and in the broader manner of expressing relationships which shall be referred to as "continuous dispersion." By continuous dispersion, we refer to an unbroken chain of amounts. For example, the traditional mode of expressing intelligence quotients has implied a *difference* between an I.Q. of 120 and an I.Q. of 121 or between an I.Q. of 69 and an I.Q. of 71. Gradually, however, clinicians have altered, by empirical necessity, their mental set toward the I.Q. and now envisage intellectual potential as continuously dispersed from "too little" to "too much" to be registered in the measuring instrument. Thus, the basic premise of this paper is that psychological phenomena are distributed in unbroken chains of altering amounts; and that our quantitative methods in psychology succeed in identifying only some of the relationships of amount within a given continuous series. A suggested inference that may follow is: Quantitative methods suffer from an apparently unavoidable fallacy in so far as the measures or scores employed represent select samples of the continuous series of the phenomenon in question.

Quantitative methods have found an extraordinary degree of application in psychology and have been insistently demanded on the American scene for a number of reasons. First of all, they represent a natural accompaniment of our mechanical age and the emphasis on engineering and physical science. Secondly, we have unwittingly attempted to emulate physics as the prototype of science without elaborating the intrinsic differences between psychology and physics. The methodology of physics makes possible a degree of detachment of subject-matter from observer which can, thus far, be attained in psychology only by doing damage to the phenomenon through artificialization. In the history of modern physics, astronomy, chemistry, etc., the recognition of the "personal equation" certainly was a boon to the development of those fields. While the facts of individual differences in perception were derived from psychology, physical scientists did not find it necessary to incorporate psychological methods, *e.g.*, introspection, along with their factual data. As psychology grew on the substrate of natural science, how-

[1] Read before a seminar at The New York Psychoanalytic Institute, March 7, 1947.

ever, not only were the facts of physics incorporated into psychology but the principal method as well: quantification.

There are other reasons to account for the quantitative emphasis in American psychology. The widely popular and growing use of mental tests with the accompanying theory of psychometrics found haven in the realm of numbers. Correlational analysis, validity and reliability determinations, etc., are cases in point. One of the most immovable stones in the whole psychometric field is the attempt to formulate criterion data in statistical terms. The feeling has also developed that psychological writings which are devoid of correlation tables are synonymous with speculation. A sense of power and security through the use of number, has engendered in some quarters an attitude of apathy or contempt for non-quantitative studies and writings. Thus the structure of psychological science has developed to a great extent upon the implicit premise that objectivity or scientific truth is obtainable only through measurement. The writer would not object to this premise if it were associated with the co-premise that psychological measurements must be rectified for variation in the measuring organism, and for the atomistic fallacy, *i.e.*, contemporary statistics generally deal with parts of phenomena taken out of their natural contexts. It is as though one would study oceanic fish life in a saline aquarium! Finally, another reason for the receptivity of American psychology toward quantification is to be found in the widespread use of "brass instrument" and physiological equipment which were forged into psychological tools.

Why are numbers thought to be intrinsically more objective than words? One cannot speak in terms of correlation coefficients or other statistical units; in the last analysis these must be reduced to words for interpretive and communicative purposes. Through the growth of general semantics we have been apprised of the limitations and essential characteristics of words and word structures. This ever-growing sophistication has not yet broken the aura which surrounds numbers. Instead of functioning merely as tools, as intermediaries between verbally formulated hypotheses and verbally formulated inferences, these means-toward-the-ends of psychological science have become ends in themselves.

We therefore come to the hypothesis that quantification need not necessarily be effected through numeration, but that the language of relativity may be at least as useful, if not more so, as numeration in our field, *e.g.*, more, less, greater than, fewer of, increase or decrease in, etc. may be utilized more profitably in formulating psychological data. For example, it has been suggested that the best criterion evolved thus far for validating psychoanalytic concepts is in terms of increasing or decreasing tensions or forces rather than through numerical quantification.

We now approach the central hypothesis of this paper. It is suggested that numeration *per se* is essentially atomistic and tends to be distortional because it leads to over- or under-determination of meaning. By over-determination we refer to the tendency to over-generalize from limited data. By under-determination of meaning we refer to an excessive narrowing of the horizon of application of specific data. For example, an I.Q. of 55 on the Wechsler-Bellevue Adult Intelligence Scale might be misinterpreted in the direction of under-determination of meaning by asserting that this merely shows that the individual has an I.Q. of 55 on the Wechsler test. Or one might commit the fallacy of over-determina-

tion by saying that this score *proves* that the individual is feebleminded. Contemporary psychological statistics would seem to be walking a tightrope between the serious pitfalls of over-determination and under-determination as two equally unfortunate end-products of the reduction of psychological phenomena to quantitative terms. It is suggested that numeration unavoidably (in the present state of statistical development) produces a state of over- and/or under-determination of meaning by means of the atomistic fallacy on the one hand (for something is measured out of context or as if there were not a context), and by inducing a false sense of security, derived ontogenetically by the use of numbers. By virtue of this sense of security one is led into inferential thinking based on the overly abstracted data of statistics.

Contemporary statistics, in becoming increasingly complicated, removes itself more and more from the world of fact which it purports to describe.

As was elucidated by Maslow,[2] the tendency has developed in contemporary psychology "to consider that the essence of science lies in its instruments, techniques, procedures, apparatus, and its methods rather than in its problems, questions, functions, or goals." It would appear as if psychological statistics is becoming more of an end in itself rather than as a means for the refined analysis of data which are obtained in the attempted solution of problems. This trend is apparent every time a psychologist is asked by his colleague, "What are the correlations?" Somehow the major unsolved issues of psychological science are frequently lost in a quasi-chaotic maze of numbers. This preoccupation with techniques rather than problems, *per se*, leads to a number of fallacies which Maslow summarizes as follows:

1. An overemphasis on polish and elegance rather than creativeness and vitality in research.
2. Giving the leading positions in psychology to technicians rather than discoverers.
3. An over-valuation of quantification for its own sake regardless of the meaningfulness of the data.
4. The forcing of problems into the molds of techniques rather than elaborating techniques for the solution of problems.
5. Intensification of a false hierarchical system among the sciences.
6. The neglect of the problems of values with a consequent blurring of the criteria for judging the worth or importance of a piece of work.

Returning to the concept of over- or under-determination of meaning, as a distortional resultant of the application of statistics to psychology, it would appear that this hypothesis is related to the more crucial question of how one goes about verifying hypotheses in psychology. The boiling cauldron of opinionated clashes between behaviorists, psychoanalysts, and Gestaltists boils down to one essential core of disagreement: What are the criteria of scientific truth in psychology, *i.e.*, what are the rules for verifying psychological hypotheses? These three major schools of psychology differ fundamentally on the advisability of applying quantitative methods toward the solution of certain problems and over the limitations and frontiers of psychological statistics. For example, many academic psychologists condemn certain areas and approaches in clinical psychology and psychoanalysis on the grounds that these fields do not subject their material to rigorous statistical checks. It

[2] A. H. Maslow, Problem-centering versus means-centering in science, *Phil. Sci.*, 1946, **13**, 326–331.

would seem to be more fruitful to ask whether the problems in clinical psychology lend themselves to solutions through the use of techniques devised in academic psychology. Much futile controversy proceeds along the lines of argument over methods, instead of a refined analysis of the problems themselves, and the forging of whatever techniques are necessary for specific problems. Psychology, like any other science, cannot hope to progress appreciably through the continued reapplication of old techniques to newer problems and by constantly working on the assumption that a quantitatively exact method necessarily yields a valid answer to any problem.

In the last analysis all quantitative material must be converted into linguistic forms. The more complicated the statistical method, the more it removes itself from the original raw data and the more speculation is involved in retracing the steps from the final statistical expressions to the meaningfulness in the original data. It must be granted that quantification facilitates regularity, systematization, and a certain kind of objectivity, but the careful use of numbers sharpens and objectifies our verbal formulations themselves to semantic analysis. In the chemical analogy, refined microanalysis is meaningless in a contaminated flask. We have passed through an era in which many statistical refinements have been elaborated but little progress in semantic insights has been made. Nevertheless statistical techniques are practically idolized, while semantics is categorized by most psychologists along with palmistry and phrenology. The broader scope of semantics, as developed thus far by Carnap, Korzybski, and others, includes basic scientific epistemology, the dynamics and social implications of scientific problems, and a vigorous onslaught in the field of verbal exactitude in the formulation of scientific hypotheses.

An undue preoccupation with quantification may also represent an escape from the central arena of psychological issues through an unstructured eclecticism and an apparent lack of emphasis upon deep creative thinking.

Many contemporary psychologists have developed a habit-pattern of viewing phenomena only in terms of their measurable units, to wit, Thorndike's famous dictum: 'Everything that exists exists in some amount. Thus, for Thorndike, every psychological phenomenon is ultimately measurable. One could reasonably take issue with this position on its *a priori* and sweeping grounds. However, even taking this formulation for what it is, the implicit essence of this dictum is the static, atomistic character of behaviorism. The writer would rather suggest that everything that exists exists in some amount with relation to everything else that exists, and that existences and amounts are relative for time-space, context, and extenuating circumstances. Thus, to express a psychological variable in quantitative terms requires a non-quantitative orientation of the measured variable to the entire context from which it was extracted; this orientation is achieved through verbal description and a verbal placement of the measured variable in time and space. If, therefore, quantitative data are to have any meaning in a pure or applied sense, they must be anchored to the original context and qualified as to limitations of meaning; this is attained through the use of words and not numbers. Words can thereby prevent numbers from yielding over- or under-determination of meaning.

From the elementalism of the ancient Greeks through the monads of Leibnitz to the stimulus-response units of Watson, a continuous thread of atomism

has been prevalent. Contemporary dynamic psychology, representing a fusion of the two major tributaries of Gestalt psychology and psychoanalysis, rests largely upon the protest against the strongly ingrained atomistic pattern of psychological thought. Statistics lends itself with tremendous facility to the atomistic approach. But how easy it is to fall into the traps of over- or under-determination! For example, a simple measure of central tendency yields over-determination of meaning by creating the illusion that a particular median describes the entire population whence it was derived. On the other hand, the median is surely something more than a mid-score in a series arranged in the order of magnitude. In a sense, a median does describe a population, but only in terms of the range, standard deviation, sampling factors, size of the population, etc. Modern statistics certainly teaches that a median has little meaning if it is not considered in relation to the general characteristics of the population whence it was obtained. These characteristics may be referred to as *contextual referants*. Furthermore, these *contextual referants* must be contrasted with those of other populations if one is to meaningfully compare medians.

While many psychologists have taken these points into account in using and examining measures of central tendency and variability, they have, in the writer's opinion, failed thus far to apply such methodological refinements to the fields of correlation and factor analysis.

Linear correlations have become a *sine qua non* in the research of academic psychology. It is very difficult to resist the temptation of subjecting two mysterious variables to a correlational procedure and harvesting a more mysterious coefficient of correlation! Now, not only is the question of relationship inadequately defined, but the correlation coefficient, in order to be rendered meaningful, must be qualified in so many ways as to make the procedure laborious and complicated. For example, the normality of the distribution of the scores of each variable is taken for granted; one should take skewness and kurtosis into account. Furthermore, the degree of communality between the two variables is represented by the square of the correlational coefficient rather than by the coefficient itself. It is also necessary to consider the essentially atomistic character of linear correlation, for one is correlating a unit with another unit, a factor with another factor, with no contextual referents. One has yet to evolve a method for correlating one Gestalt with another Gestalt, or for incorporating time-space referants into a correlational coefficient. Of course, the saddest blind spot of all is the construction and supposed verification of psychological hypotheses through the use of correlational analysis ignoring the fact that correlations give measures of concomitance and not causality. Thus the coincidental fallacy, a relic of primitive and common-sense thinking, is fostered by correlational methods which are all right in themselves but have been misused and overused in our field more than in any other discipline. It it also clear that much more must be done in the field of multiple correlation and regression equations and that psychologists become less dependent upon linear correlation in proportion as they recognize that this procedure is a natural accompaniment of atomism and reductionism and that correlational methods were developed in the fields of agronomy and population statistics. This exemplifies a point emphasized above that the difficulty with psychological statistics has been the presaging of hypothesis formation by avail-

able methods. In other words, since correlational methods were carried into the field of psychology, the modern psychologist is faced with two alternatives in this regard: One is to go through the qualifying procedures in the use of correlation coefficients, as suggested above; this becomes more laborious than a careful detailed verbal delineation. The second alternative is to indulge the blind use of correlational methods and, thus, satisfy one's inner needs for being "scientific." The psychological journals are saturated with an abundance of papers which follow the second alternative. The discerning reader is continuously asking himself the inevitable "so what?" after reading much of this material. For, after all, if a paper starts out with a problem posited in verbal terms and winds up with a summary of quantitative results, the job is only half done and the reader remains in mid-air. On the other hand, the careful, configurational and dynamic *interpretation* of the statistical results is extremely difficult in the light of the intrinsic limitations of psychological statistics thus far. Would it not, therefore, seem more plausible and necessary for us to elaborate rules and techniques for global interpretation of statistical results rather than attempting to refine the statistical methods, *per se?*

A frontal and unashamed self-examination of psychology would, it is suggested, reveal an increasing trend toward the substitution of the easy and automatic statistical methods for deep creative insights and formulations. If creative inspiration is being plowed under by a rigid insistence upon numeration, then there is little cause for surprise at the paucity of fundamental psychological discoveries. The house of contemporary psychology is a veritable confusion of development and refinement of methods—but methods for what problems?

The epics of science, the discoveries of lasting and stimulating value, have carried in their wake the elaboration of new methods. Few significant discoveries have followed the persistent reapplication of old methods which were applicable to older, now solved, problems. The discovery of the tubercle bacillus opened a large chapter in bacterial differential staining. The discovery of X-rays preceded the application of this remarkable phenomenon to medical diagnosis. Roentgen did not apply traditional methods to the discovery of the X-ray, nor did he believe that he would benefit clinical medicine. His discovery resulted from the simultaneous occurrence of a phenomenon not explicable in terms of established theory, together with his keenness of perception and reasoning. Roentgen had the flexibility of mind to formulate new hypotheses for old data and to pick up the small thread of a chance observation and carry on from that point. While we certainly know very little about the psychology and psychopathology of creative thinking, one thing seems clear: Flexibility is the life blood of any science, especially one so encompassing as psychology; and that a method is valuable only to the extent that it will facilitate the elaboration of hypotheses which cannot be adequately approached by that method. Therefore the usefulness of a scientific method inheres in the explicitness of the limitations of its usefulness. In psychology the use of statistics could become more meaningful should psychologists make its limitations more explicit.

It might be helpful to illustrate the presaging of research methods by carefully refined formulations of problems and hypotheses from the field of psychological science itself. The persistent efforts of the Würzburg school, headed by Mach, to discover something about the psychology of thinking

through the repeated use of the method of controlled introspection yielded little tangible result; the method of hypnosis as used in the psychiatric clinic in Paris proved inadequate to account for and treat certain kinds of neurotic disorders. Freud, essentially a pragmatist, formulated certain hypotheses out of which the psychoanalytic method grew and out of which, in turn, an elaboration of psychoanalytic theory ensued. Thus it is seen that vigorous advances in psychology, as in other sciences, follow upon a reversible interaction of methodological refinement and formulation of hypotheses. The attempt to force a new hypothesis out of an old method is essentially inductive; the effort to formulate a new hypothesis *purely* by inspiration and then draw whatever strands of new methods one can out of this hypothesis is essentially deductive. All will agree with modern logicians that science is inductive-deductive, and to preoccupy one's self with either one of these two components of scientific reasoning is rigid and misleading. Therefore quantification in psychology helps, to a certain extent, to lay the ground-work for new formulations in a given area, provided that area lends itself to measurement in our present stage of development. This tendency has, it is suggested, been overdeveloped on our contemporary scene with a derogation and lack of emphasis upon clearly formulated and meaningful problems, the semantic refinement of their manner of expression, and a *subsequent* delineation of the role of statistics in the attempted solution of the particular problem.

Lest the reader deduce that a polarization of quantitative versus non-quantitative methods is posited here, it should be stressed that among the fascinating revelations of scientific history one finds the greatest progress and most rapid rate of growth in borderland fields such as physical chemistry, biochemistry, psychosomatic medicine, etc. That is to say, two possible borderlands in psychological science which remain to be explored are:

1. The dynamic inter-connections between the various quantitative and non-quantitative methods with an elaboration of the mutual dependencies and limitations of each.
2. The study of the relationships between numerical and semantic analysis toward the end that the present dichotomy between measurement and description be permanently relegated to the past of Aristotelian logic. This would go a long way toward the reduction of tension and hostility between academic psychology, based largely upon numeration, and clinical psychology, accustomed essentially to verbal characterizations.

This brings to the fore the essential basis for clash between academic and clinical psychologists, since the primary issue at stake is quantification. The academic psychologist has restricted the scope of his investigations in order to fit them into available statistical methods. This is only half of the methodological fallacy involved, for the academic research worker takes too much for granted in the utilization of present day statistics. A fusion should be attempted in the form of discovering the area of communality between the strengths of quantification and the weaknesses of verbalization, and *vice versa*. In this manner, these two great tools of the human mind, word and number, would buttress and facilitate each other at their mutually weak points rather than thwart the development of psychology by mutual inhibition. Since both word and number are symbolic representations of inner and outer reality and constitute our chief means of social communication, it might be well for research workers to study the methods of hypothesis for-

mation and hypothesis verification in psychology as a primary pursuit. The relative emphases and limitations of words and numbers would then grow naturally out of the rules derivable therefrom. Only in this way can psychological research gain a certain vitality and courage which it lacks as long as it continues to build its hypotheses on the structure of old methods.

Not only does the use of statistics tend to produce over- or under-determination of meaning, but also over or under self-confidence of the psychologist. On the one hand, the psychologist does not trust himself to resort to sufficient use of imagination and intuition. On the other hand, quite paradoxically, he appears to trust himself too much when he fails to realize that the use of numbers does not preclude the operation of bias and the selective nature of human attention and perception, let alone the powerful influence of unconscious motivation. Nor does the use of numbers necessarily imply objectivity and authenticity; a cloud of statistics may very well befog real issues.

Those who tend almost to idolize quantitative methods function in their work as if the following were synonyms: *measured, objective, scientific, empirical, experimental;* and as though the following, likewise, were identities: *non-quantitative, speculative, Freudian, mystical, clinical, unscientific.* Both academic and clinical psychologists hurl the epithet at each other of either being unscientific through failure to adequately use numbers or for being unscientific for failure to work with meaningful problems. It is surely time for a fusion of emphases as between quantification for its own sake and the neglect of the positive virtues of statistics. This fusion can be partly attained, it is suggested, through the use of ordinal scales and semantics.

In the field of psychometrics and aptitude testing it is often forgotten that a score of zero on a test does not necessarily mean zero ability or aptitude; *a score is not a fact.* The score has been abstracted from other possible measures of the same ability; it is far from perfect in validity and reliability, etc. These limitations of test scores and their de-contextual character have been worked out by psychometricians and theorists. But the implication is that the gulf between data and facts is ever-widened by each further operation of a statistical nature. Hence, it would seem desirable to maintain simpler statistical procedures. Measures of central tendency and variability and measures of the significance of the difference in means, etc., remove themselves only slightly from the original data and the possible interpretations of these data. On the other hand, correlational and factorial analyses depart very far indeed from the original data and make many more assumptions which are not entered explicitly into the working system. One test which might be applied for the safety of statistical method is the relative ease with which it can be explained to the intelligently unsophisticated.

All statistics thus far have been developed in the direction of inter-individual differences. In conjunction with such clinical concepts as "normality" and "self-consistency," it may be interesting to develop intra-individual statistics. However, this does not imply the carry-over of inter-individual methods to the problems of intra-individual variation but rather the elaboration of new methods. To be able to measure the amount of erratic tendency in an individual or the variation in his instinctual strivings would find valuable

application. Perhaps inverted factor analysis will supply the answer. It must, however, be based upon some formulation of personality theory which has a clinical anchorage.

In conclusion, quantification facilitates the growth of psychology to the extent that it is subsumed under the broader realm of qualitative description and explanation. Statistical methods inhibit vigorous growth in psychology to the extent that they promote atomistic, categorical thinking, and over- or under-determination of meaning. Quantitative emphases are further damaging in that they imply that qualitative methods are not subject to the same rules of verification and authenticity, and are necessarily less reliable than quantitative methods.

The expansion of the frontiers of psychological science will surely be facilitated by some fusion of qualitative and quantitative approaches through a new, yet-to-be-devised, system of semantic-statistical analysis.

[MS. received May 16, 1949]

A METHOD OF GRADED DICHOTOMIES FOR THE SCALING OF JUDGMENTS [1]

BY FRED ATTNEAVE

University of Mississippi

I

The methods based upon Thurstone's Law of Comparative Judgment (3, 4) seem the most rational yet devised for ordering a group of stimuli or concepts on a psychological continuum when the corresponding physical continuum is unknown, unmeasurable, or otherwise lacking. These methods suffer from certain experimental handicaps, however, which seriously limit their practical usefulness. Thus, if all possible paired comparisons are made, the experimental session must include $\frac{n(n-1)}{2}$ judgments: even with a fairly small number of stimuli (*e.g.*, 15) the task is a tiring one, and with a large n (*e.g.*, 50) it becomes virtually impossible. Considerable economy is achieved by requiring subjects merely to rank-order the stimuli, and deriving the necessary paired comparisons from the data thus obtained, but this method (besides involving an extremely laborious tabulating procedure) usually precludes the possibility of obtaining judgments from subjects in groups, since each subject must be given the experimental materials to manipulate.

Rating-scale methods, in contrast, represent an ideal of experimental convenience: the minimum number of judgments involved is only one for each stimulus, and the stimuli may usually be displayed to large groups of subjects simultaneously. Unfortunately, however, the rating-scale yields seriously distorted results when scored according to the obvious assumption that the scale intervals are psychologically equal. Hevner (2), on comparing empirically the equal-appearing intervals method with the Thurstone methods, found a markedly curvilinear relationship between the results: the intervals near the end of the scale were found to be "larger" than those in the middle. The source of this distortion is fairly evident on *a priori* grounds. The ratings of a given item by a number of judges will, when unrestricted, tend to distribute themselves normally about the hypothetically "true" rating; but when the "true" rating falls at an extreme, variability is possible in only one direction. The distribution obtained in the latter case will therefore be skewed away from the end of the scale, yielding a mean (or median) too close to the center.

The method about to be described preserves the experimental simplicity of the rating-scale technique, but treats the ratings in a manner which rationally avoids the end-distortion just mentioned.

II

This method may be considered an extension of the Method of Single Stimuli (5, 6). In the latter the subject is required to classify each of a number of stimuli (*e.g.* lifted weights) into one of two opposite categories (*e.g.* "Heavy" and "Light"), an intermediate judgment being also permitted

[1] The author wishes to thank Dr. Quinn McNemar and Dr. Donald W. Taylor, both of Stanford University, and Dr. Harry Helson, of Brooklyn College, for reading the initial draft of this article.

in some cases. After becoming familiar with the series, the subject adopts a fairly stable subjective standard, or adaptation-level (1), so that the frequencies with which successive stimuli are assigned to a given category describe a normal ogive.[2] Knowing the proportion of times which a subject (or group) assigns a given stimulus to each of the categories, we may consider that the psychological value of that stimulus is given by the deviate of the normal curve corresponding to that proportion. If, for example, four different weights have been judged "Heavy" 10, 50, 60, and 80 per cent of the time, respectively, we find the corresponding deviates -1.28, zero, $+.25$, and $+.84$, and assume that these figures indicate the spacing of the stimuli along the psychological continuum, as illustrated in Fig. 1.

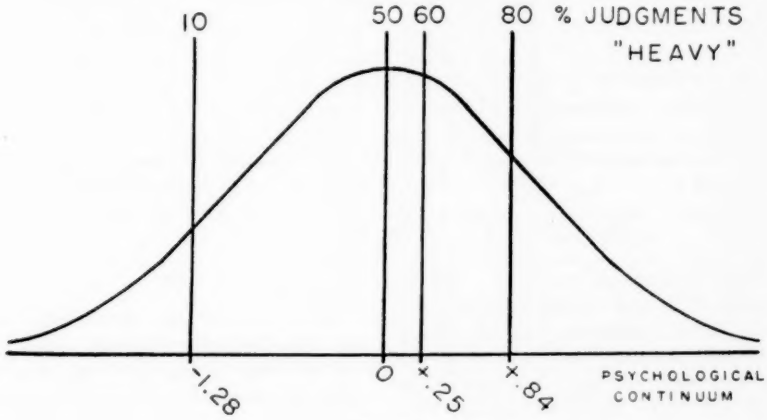

FIG. 1

Let us now consider a situation in which three instead of two categories of judgment are permitted: these might be "Light," "Medium Heavy," and "Heavy" (for illustrative purposes we choose a middle category which is not quite neutral in its connotation). We may dichotomize the ratings thus obtained for a given weight about either of two points: the "Light"–"Medium Heavy" division or the "Medium Heavy"–"Heavy" division; that is to say, we may combine the "Medium Heavy" judgments with either the "Heavies" or the "Lights," and thence proceed as in the two-category case. Ideally these two ways of dichotomizing should give us precisely equivalent results in terms of the obtained *separations* of stimuli on the psychological continuum. Thus the four weights considered above might be ordered as in Fig. 2. In this illustration the psychological values (X_1 and X_2) obtained from the two ways of dichotomizing differ by 1.25 σ units. This constant represents the psychological width of the "Medium Heavy" category. In a practical situation, where the correspondence is not so precise, the mean difference between X_1 and X_2 would be taken as the width of the intervening interval.

[2] We are here making certain assumptions—essentially those of Thurstone's Case V (4)—which we shall presently examine in more detail.

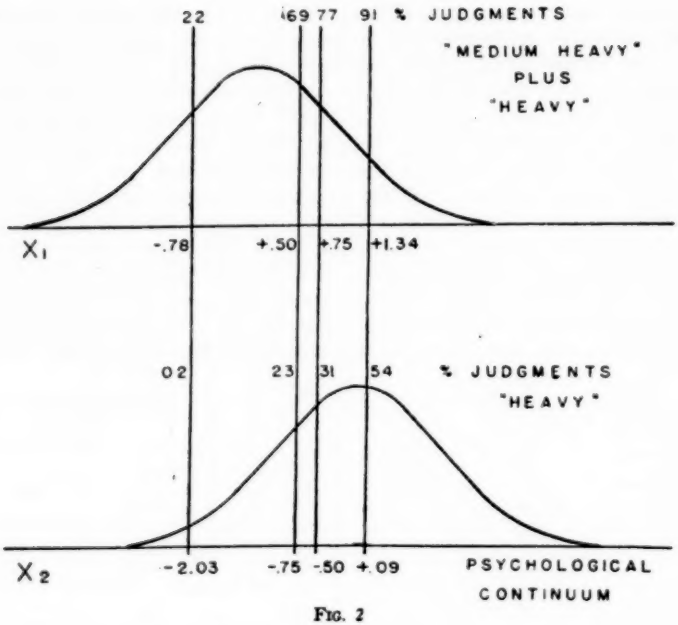

FIG. 2

Such reasoning may be applied equally well to a rating scale of more than three categories. If we use a scale with k categories, there are $k - 1$ ways in which we may dichotomize the obtained data. Since the ratings of a given stimulus will usually not cover the entire scale, the relative psychological position of that stimulus will be indicated only by those dichotomies which actually divide its ratings, the X-values corresponding to 0 per cent and 100 per cent being indeterminate. The size of the interval separating adjacent dichotomies may be inferred from the X-values of stimuli common to both, as suggested above; thus, by the addition of suitable constants, X-values from all the dichotomies may be rendered comparable, and fitted together on a single psychological continuum. We shall refer to such a treatment of ratings as the Method of Graded Dichotomies.

III

Application of the method involves the following computational steps:

1. A table is prepared, the rows of which represent stimuli and the columns rating-scale categories, showing for each stimulus the percentage of ratings falling within each category.

2. From this a second table is made, with a column for each place at which the ratings are to be dichotomized, showing for each stimulus the percentage of ratings falling above each cutting-point. Most complete use of the data is presumably made by using all possible $(k - 1)$ dichotomies, but satisfactory results may be obtained with some omissions if k is large and ratings are broadly dispersed.

3. By the use of a table of the normal curve (or, preferably, of the normal ogive) the above percentages are translated into x/σ units. These X-

values, as we shall call them, are designated as positive when the percentage above the cutting-point is greater than 50 per cent, and negative when it is less than 50 per cent. Extreme percentages (usually those less than 5 per cent or greater than 95 per cent) are discarded.

4. For each pair of adjacent dichotomies, the mean difference between the X-values of those stimuli represented in both is calculated. This is taken as the magnitude of the interval separating the two cuts.

5. To the X-values of each dichotomy is added the sum of the intervals below that dichotomy. An additional constant may be added to all the X-values in order to make the smallest a positive number. The purpose of this operation is to refer all the X-values to a common (arbitrary) origin, so that they will be comparable with one another.

(The figures obtained in steps 2, 3, and 5 may be recorded in the same table if sufficient space is allowed in each cell.)

6. Each stimulus now has one or more (equated) X-values, the mean of which is taken as the best estimate of its relative position on the psychological continuum. In theory the X-values for a given stimulus should be identical: if they differ more than considerations of experimental error would lead us to anticipate, we may suspect that certain assumptions of our method are incorrect. This contingency will be considered later.

IV

The ratings obtained by Hevner (2) from 370 subjects on the goodness of 20 specimens of handwriting have been scaled by the Graded Dichotomies Method in the manner just described. Hevner's data constitute excellent material for the empirical validation of this method, since she also applied two different methods based upon Thurstone's Law of Comparative Judgment (*Paired Comparisons* and *Order of Merit*) to the same specimens, using the same subjects. Her rating scale contained eleven categories, making possible ten dichotomies, but the dispersion of the ratings was such that only six of the ten were used. Cuts were made between the following categories: 2 & 3, 4 & 5, 6 & 7, 8 & 9, 9 & 10, 10 & 11.

Table I shows the value obtained for each specimen by the present method and by the three methods which Hevner employed. In Figs. 3, 4, and 5 the values obtained by the Graded Dichotomies procedure are plotted against those from each of the earlier methods. The

TABLE I

HANDWRITING SCALE VALUES OBTAINED BY FOUR METHODS

Specimen	Method			
	Graded Dichotomies	Order of Merit	Paired Comparison	Eq.-appearing Intervals
b5	.00	.00	.00	1.60
g3	.44	.45	.44	2.15
h3	.66	.65	.45	2.80
y5	.71	.85	.47	2.80
f4	.40	.62	.55	2.25
k2	.21	.43	.67	1.80
t4	.49	.80	1.43	2.35
f2	1.66	1.94	1.72	5.13
n2	2.17	2.66	1.88	6.23
b4	2.05	2.88	2.44	6.10
t3	3.18	4.26	3.20	8.10
z3	2.87	3.79	3.30	7.70
j2	3.15	4.22	3.52	8.10
x3	3.51	4.90	4.33	8.70
r2	3.81	5.44	4.44	9.30
u3	4.75	6.85	5.55	9.95
a4	4.71	6.77	5.53	9.90
g4	4.79	7.39	5.87	10.15
j3	5.60	8.27	7.43	10.40
d2	5.92	8.58	7.69	10.50

The above data, with the exception of the Graded Dichotomies values, are reproduced from Hevner (2).

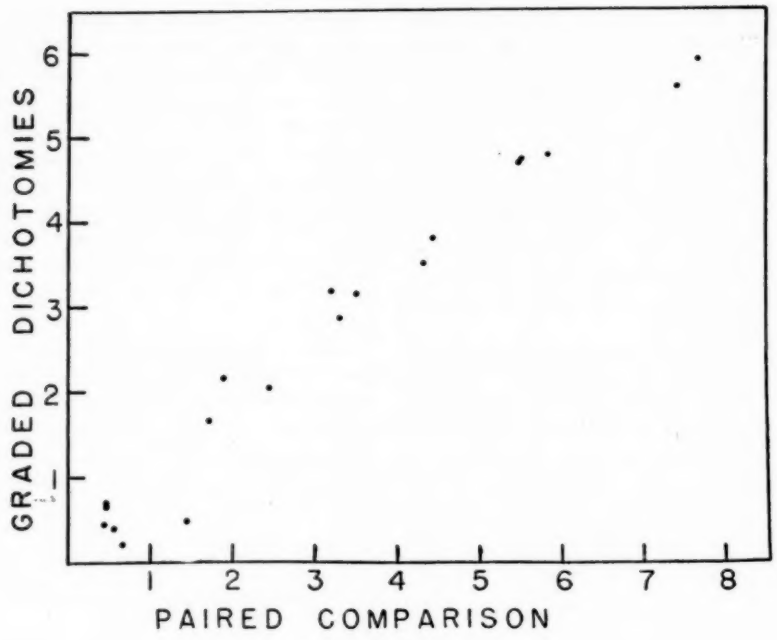

FIG. 3

agreement between the present method and the two Thurstone methods (Figs. 3 and 4) is quite satisfactory, being approximately as good as that of the Thurstone methods with each other. Whereas these two plots show little suggestion of non-linearity, the relationship between the Graded Dichotomies values and the Equal-Appearing Intervals values (Fig. 5) is conspicuously curvilinear, as we should expect from Hevner's findings. It should be pointed out that the latter two sets of results are based upon precisely the same raw data, so that discrepancies are entirely a function of differences in treatment.

V

Some consideration of the assumptions involved in the Method of Graded Dichotomies, and of its relationship to the Comparative Judgment methods of Thurstone, seems desirable at this point. In the first place it may be emphasized that the method does not assume subjective equality of rating-scale intervals, but instead infers the relative magnitudes of these intervals from the rating behavior of the subjects. Neither does it assume that the stimuli are distributed normally, or in any other specific way, with respect to the attribute in question. It does assume (1) that ratings of a given stimulus will, in so far as possible under the restrictions of the scale, be distributed normally about a hypothetically "true" rating, and (2) that the distributions of ratings for the various stimuli will have identical standard deviations except in so far as the distributions are subject to curtailment, this constant standard deviation being the unit in which ratings are finally scaled.

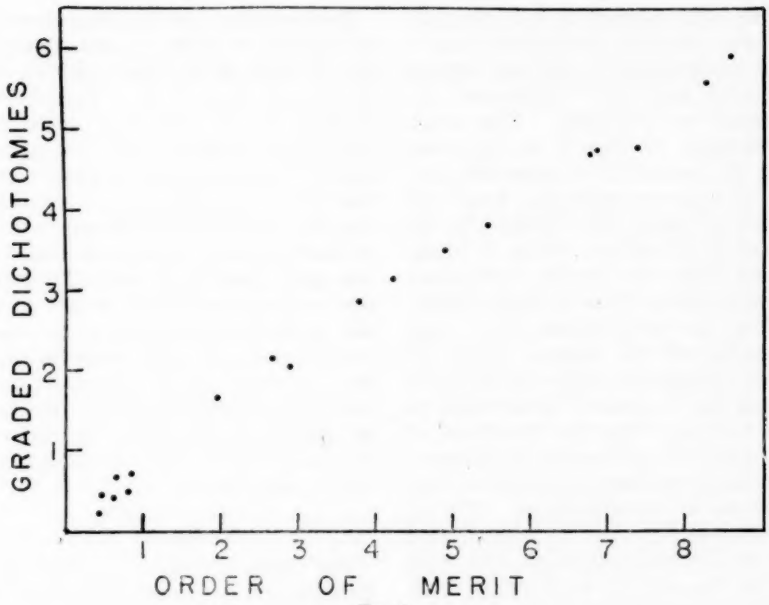
GRADED DICHOTOMIES
1
2
3
4
5
6
1
2
3
4
5
6
7
8
ORDER OF MERIT

FIG. 4

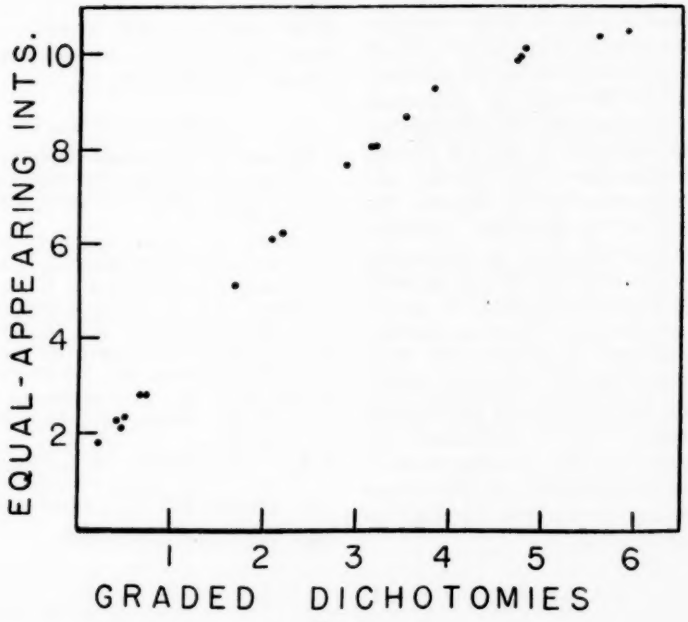
EQUAL-APPEARING INTS.
2
4
6
8
10
1
2
3
4
5
6
GRADED DICHOTOMIES

FIG. 5

The latter assumption is in principle the same as that of Thurstone's Case V (4), which demands that the stimuli compared have equal "discriminal dispersions," or ambiguity. (The other requirement of Case V is concerned with the correlation of judgments, and has no apparent application here.) It should be noted that Hevner (2) assumed the conditions of Case V in applying both the Paired Comparison Method and the Order of Merit Method.

If we are led to believe, either from inspection of the original ratings or from unexpectedly large discrepancies among the X-values to be averaged in the final step, that the dispersions of ratings for the various stimuli are genuinely unequal, some adjustment for this condition is clearly necessary. The desired correction may be approximated by the following operations: (1) After completion of step 4 above, for each stimulus the difference between the extreme unequated X-values (*i.e.*, the difference between the X-values from the highest and lowest dichotomies in which the stimulus is represented) is obtained. The sum of the intervals (computed in step 4) separating the dichotomies which provide these X-values is divided by this difference. The quotient obtained represents a ratio between the dispersion of a particular stimulus and the mean dispersion of comparable stimuli (it is necessary to bear in mind that the relative magnitudes of the X-values are *inversely* related to the dispersion of the stimulus). (2) This quotient constitutes a correction factor, by which each X-value of the stimulus under consideration is multiplied. By the present operation the dispersions of the various stimuli are equated. Steps 5 and 6 may now be followed using the corrected X-values.

The procedure just described extends the application of the Graded Dichotomies Method to situations covered by Thurstone's Cases I–IV. Two disadvantages of the procedure must be pointed out, however: (1) It assumes that the *mean* dispersion of stimuli represented in one dichotomy is equal to the mean dispersion of those represented in another; this assumption may be only approximately correct. (2) Situations may arise in which certain (usually extreme) stimuli are cut by only one dichotomy, in which case the data are insufficient for the estimation and correction of their dispersions. The latter difficulty is minimized by the use of a scale with a relatively large number of categories.

SUMMARY

The treatment of rating-scale data by a Method of Graded Dichotomies is described. This method avoids the assumption of equal scale intervals, and yields results closely comparable to those obtained by Thurstone's paired comparison techniques.

BIBLIOGRAPHY

1. HELSON, H. Adaptation-level as frame of reference for prediction of psychophysical data. *Amer. J. Psychol.*, 1947, **60**, 1–29.
2. HEVNER, K. An empirical study of three psychophysical methods. *J. gen. Psychol.*, 1930, **4**, 191–212.
3. THURSTONE, L. L. Psychophysical analysis. *Amer. J. Psychol.*, 1927, **38**, 368–389.
4. ——. A law of comparative judgment. PSYCHOL. REV., 1927, **34**, 273–286.
5. VOLKMANN, J. The method of single stimuli. *Amer. J. Psychol.*, 1932, **44**, 808–809.
6. WEVER, E. G., & ZENER, K. E. The method of absolute judgment in psychophysics. PSYCHOL. REV., 1928, **35**, 466–493.

[MS. received May 16, 1949]

AN APPLICATION OF ${}_sE_R$ QUANTIFICATION PROCEDURE

BY DAVID ZEAMAN

University of Connecticut [1]

Introduction

A methodology has been proposed by Hull, Felsinger, Gladstone and Yamaguchi (6) for the quantification of the construct, reaction potential (${}_sE_R$). In their paper (p. 252), they state:

"As has been pointed out, there are many different indications of reaction potential. It is probable that some forms of learning can be found which yield two or more such objective indications of the very same habit strength acquisition, each one of which could be used independently as reaction latency was above. It would seem that such curves of learning even though secured from quite different data should agree as to the type of best fitting equation and also, within the limits of the data sampled, as to the various constants. Such a determination may very well prove to be a critical test of the methodology."

It is the purpose of the present paper to apply the ${}_sE_R$ quantification procedure to two different measures of performance in the same learning situation.

The Data

Empirical learning curves for the two objective measures are presented in Fig. 1. The two curves show the course of change of median latencies and median running-times of 40 rats during 18 conditioned reactions on a simple elevated straight runway (Graham-Gagné type). Latency is the time the animal takes to leave the starting box once the door is opened; running-time, the period required for traversal of the three-foot runway to the goal box.[2]

The more important conditions of reinforcement include: (1) temporal factors—only one reinforcement trial a day was given with no delay of reinforcement; (2) kind of reinforcement—"process" cheese; (3) amount of reinforcement—an intermediate amount in the appetite range, 0.60 gram on each trial. The radical spacing of trials is necessary to eliminate inhibition factors which would complicate the quantification of what is assumed to be a single factor of excitation.

As measures of excitation or reaction potential, both latency and running-time have been shown to be sensitive to changes in the common experimental variables and operations; that is, they have been used to demonstrate many of the basic behavioral processes (1, 2, 4, 7, 8, 12, 14). In Hull's system (5) response latency appears formally defined as one of the four observable anchors of ${}_sE_R$. Running-time, however, is not given a formal status although it is used by Hull (5) a number of times as an index of reaction potential. As has been pointed out (14), running-time could easily be included within the framework set down in the *Principles of Behavior* by extending the category (A) amplitude of re-

[1] This paper was written while the author was a staff member of the Department of Psychology, Brown University.

[2] The details of the apparatus and procedure have been described in a previous study concerned with the effects on latency of varying amounts of reinforcement (14). The data of the present paper are also derived from the experimentation of that study, performed in the Psychological Laboratory of Columbia University.

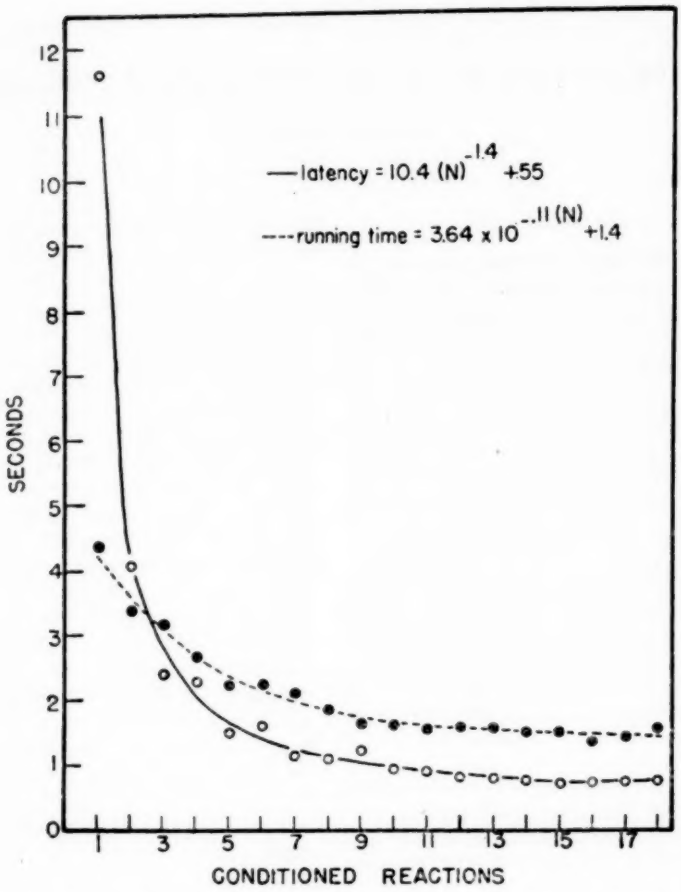

FIG. 1. Two empirical learning curves representing the performance of forty rats in a simple runway situation. Both median latencies and median runway-times are plotted as a function of the number of conditioned reactions. Different equations are shown fitted to the two sets of data.

sponse to include striate muscle responses as well, or by making it a special case of a latency measure.

The empirical learning curves for the two measures are quite different, as can be seen in Fig. 1. The range of scores is much greater for latency than for running-time, and would be greater yet if the first trial (counted as unconditioned) were included. The apparent asymptote of each curve is clearly different, the latency measure yielding a greater final strength. The two curves are similar in that they are both fairly regular, negatively accelerated, decreasing functions, undergoing the major fraction of their total change in a few trials, and approaching an asymptote within the 19 trials given. Before the ${}_{S}E_{R}$ quantification procedure is ap-

plied to these two different measures to discover if, indeed, a common function results, it is desirable to express with more precision these similarities or differences between the two unquantified, empirical functions. To this end, curve-fitting procedures were introduced to give mathematical expression to the two learning curves.

CURVE-FITTING

There are of course an infinite number of equations which can be fit to any set of empirical data. Ordinarily in the absence of a theoretical prediction of equation form, heavy use is made of the Principle of Parsimony in trying first, and accepting the best fit among, the simpler equation types.

In selecting equations to be fitted to the two empirical learning curves in Fig. 1 it is possible to be guided not only by the general shape of the curves themselves, but by some rational considerations derived from the nature of the phenomenon under investigation, and by previous attempts at such curve fitting to related data. The rational considerations include at least these: (1) the equation form must yield a finite expression for the starting point or first conditioned reaction; (2) it must yield an expression for a finite asymptote; (3) it must not generate any negative values.

Among the simpler equation forms which generate appropriately-shaped curves, and meet the criteria, these were tried:

(1) a power function, $Y = aN^{-b} + c$,
(2) an exponential function, $Y = ae^{-bN} + c$,
(3) a log power function, $\log Y = aN^{-b} + c$,
(4) a log-exponential function, $\log Y = ae^{-bN} + c$,

where Y is the measured performance and N the number of conditioned reactions. Felsinger, Gladstone, Yamaguchi, and Hull (2) have found that a power function closely describes their latency learning curve. Graham and Gagné (4) used an exponential type of equation with success in predicting the course of acquisition with latency measures. For statistical purposes, logarithmic transformations of time measures are frequently employed, so the goodness of fit of the power and exponential function should be tested with latency and running-time expressed in logarithmic form.

A graphical technique tested each of these equation forms for appropriateness and goodness of fit to the empirical data. For the power function this test consists of plotting on log-log graph paper the difference between the final limit, or asymptote of performance, and the actual performance (median latency of the group) on each trial against the number of previous reinforcements.[3] The rectilinearity of such a plot confirms the choice of equation form and measures the goodness of fit. A difficulty arises, however. Since the asymptote of performance was neither known nor easily estimable from a plot of the raw data, it was necessary to assume various values and test for linearity of the log-log plot. An assumed asymptote of .55 second yielded a curve most closely approximating a straight line. The slope of this line was -1.40, the Y-intercept 10.4. These two values are the remaining constants in the power function which was sought. The equation then reads:

$$\text{latency} = 10.4\ N^{-1.40} + .55.$$

Figure 1 shows this equation to represent adequately the latency data.

[3] Accepting Hull's convention, the first trial is regarded as a measure of unconditioned strength, and numbering of performance measures starts with the first conditioned reaction—on the second day's trial.

The running-time data were first tested with the power function since that has already proven applicable to the latency data. The most nearly linear of the log-log plots (as before) resulted when an asymptote of zero was assumed. Therefore this equation form was ruled out on the rational grounds of impossibility of a zero asymptote.

Next, the running-time data were put to the graphical test for the exponential equation form. The test requires a semi-log plot; the logarithm of the difference between actual running-time and an assumed asymptote plotted against number of previous reinforcements. The results were most satisfactory with an asymptote assumed at 1.4 seconds. Computation by the method of averages gave the slope and Y-intercept of the straight-line function, which in turn are the remaining two constants in the final equation:

$$\text{running-time} = 3.64 \times 10^{-.11\,N} + 1.40.$$

This same test for the exponential type of equation was applied to the latency values, to the logarithm of the latencies, and to the logarithm of the running-times. The log scores of each measure were also tested with the power function. The results of these five tests were negative. Either there were obvious non-linear trends in the log-log or semi-log plots when a reasonable range of asymptotes had been sampled, or the random deviations (squared) from the best fitting straight line averaged out larger than those found in the tests yielding the two equations written above.

${}_sE_R$ QUANTIFICATION

This section prepares the answer to the question: What does ${}_sE_R$ quantification do to the two different empirical learning curves—for latency and running-time? The details of the methodology have been set forth with admirable simplicity by H.F.G. & Y.[4] as adapted from Thurstone (9, 10, 11). It is only necessary therefore to present evidence of the soundness of the assumptions on which the methodology is based before giving the results.

The first assumption concerns the homogeneity of the variability in σ_k, an estimate of the dispersion of the individual variables, said to be related to Hull's "oscillation" of reaction potential. The σ_k's are used to rectify certain scale distortions in the quantification. The differences between any pair are assumed to be not very large. A two-to-one ratio of largest σ_k to smallest σ_k is considered sufficient homogeneity. Figure 2 shows the trial-by-trial changes in σ_k. It can be seen that as learning proceeds there is much variability in the estimated individual dispersions: random variation due to limited size of the sample, and systematic variation due to the progress in learning. These curves resemble the σ_k function presented by H.F.G. & Y. in that they have a wave-like appearance. They differ in that the waves are greater in amplitude and perhaps shorter in length. The increase in amplitude of variation would be expected since the N of this study was only two-thirds that of the H.F.G. & Y. study.

There are two points of major importance to be made about Fig. 2. The first involves a comparison between the order of magnitude of the σ_k variation in this and in the H.F.G. & Y. study; the second considers the relation between σ_k variability of the latencies and of the running times. Inspection of Fig. 2 shows that most of the σ_k's of the present study fall within the range of those in the H.F.G. & Y. study.

[4] For convenience, continued reference will be made to the series of papers by Hull, Felsinger, Gladstone and Yamaguchi (2, 3, 6, 13) by the abbreviation H.F.G. & Y.

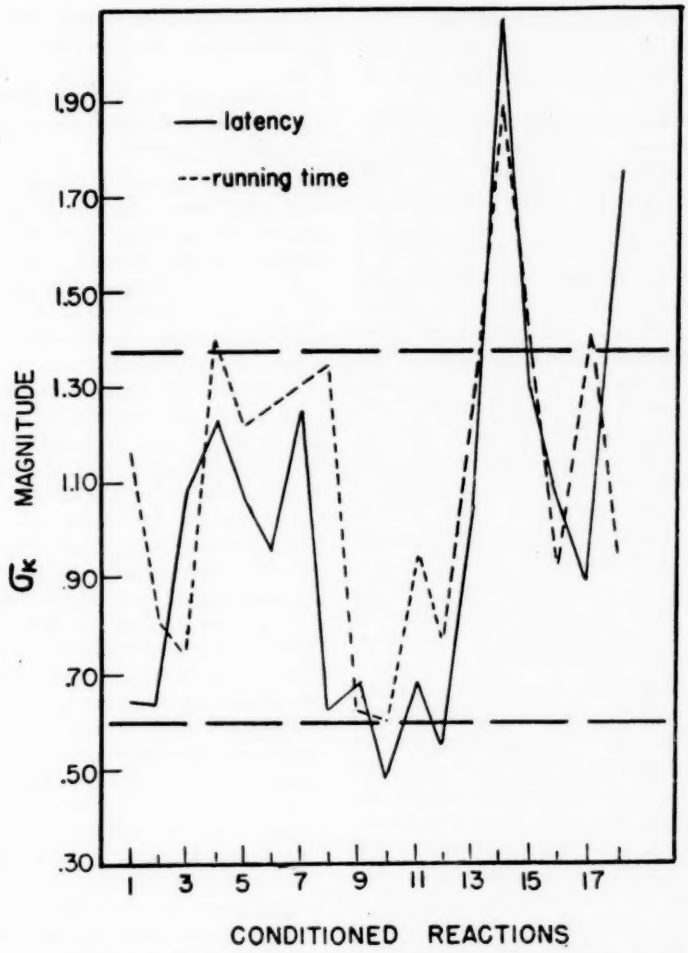

FIG. 2. The dispersion of the individual variables (σ_k), used in the present quantification procedure, as related to the progress of learning.

The two horizontal dashed lines indicate the range within which fell all the σ_k values used in the H.F.G. & Y. study (6).

If three of the 36 σ_k values were excluded, the remaining σ_k's would approximate the 2–1 ratio criterion. This degree of heterogeneity throws some doubt upon the assumption of homogeneity, but it is not considered great enough to completely vitiate the ultimate comparison of the ${}_sE_R$ functions generated by the quantification procedure. The close parallelism of the σ_k curves for the two measures fortifies this judgment since it is likely that whatever distortion is produced by the heterogeneity would affect both ${}_sE_R$ functions roughly to the same degree leaving the relation between them un-

disturbed. The primary purpose of the paper is of course to find whether the resultant ${}_sE_R$ curves are the same or different.

A second assumption upon which the quantification procedure is based is tested by plotting against each other any two columns of the tables enumerated by H.F.G. & Y. 5 and 7. Ten pairs of columns from Table 5 for each of the two measures were plotted, and ten pairs of columns for each measure from Table 7. The outcome of these tests was substantially the same as that found by H.F.G. & Y. (6). The plots from Table 5 were approximately linear, as were those from Table 7, but the slopes of the Table 7 plots came closest to the desired unity. The slope of each graph was computed. The average slope of the 10 plots for the running-time data from Table 7 was .98 with an average deviation of .09. The average slope for the corresponding latency values was .97, with an average deviation of .09.

These findings confirm the appropriateness of Thurstone s Case V method for these data by lending validity to the dispersional assumptions underlying it.

The quantification is based upon the assumption of normality of distribution of ${}_sE_R$'s and ${}_sE_R$ differences. The soundness of this assumption is given support by the evidence presented in the last paragraph. A more precise method of testing the normality of ${}_sE_R$'s has been employed by H.F.G. & Y. The test consists of:

1. completing the quantification using the normal probability tables and deriving the equation which expresses the relation between ${}_sE_R$ and ${}_st_R$;
2. converting a number of asymptotic latencies of each animal into ${}_sE_R$'s by use of this equation;
3. constructing a frequency distribution of these ${}_sE_R$'s corrected for individual asymptote differences;
4. making another distribution of the ${}_sE_R$ differences at the latency asymptote;
5. computing the deviation from normality of these distributions in terms of skewness and kurtosis.

H.F.G. & Y. applied this test to their data and found the distributions of ${}_sE_R$'s to deviate from normality in the direction of leptokurtosis. They corrected for this distortion by constructing an empirical leptokurtic probability table and running through the quantification procedure again.

This test and refinement were not employed in the present study for several reasons. The prohibitive amount of work involved in the computations necessary for the re-quantification may be, perhaps, a real reason. Good reasons include the fact that only about five asymptotic latencies were available for each animal as compared with the 20 such latencies in the H.F.G. & Y. study. This would severely limit the size of the distribution of ${}_sE_R$'s and thus limit the accuracy of the correction to be made on the basis of the shape of the distribution. The best reason lies in the fact that H.F.G. & Y. found the correction for leptokurtosis to make so little difference in the final ${}_sE_R$ equation. Only one constant, was changed and that by a factor of only 10.6 per cent.

There seem to be sufficient grounds for the assumption that whatever distortion is caused in the two ${}_sE_R$ curves of this study by the use of normal probability tables is not great and/or does not affect the two ${}_sE_R$ curves differentially.

RESULTS

The two curves presented in Fig. 3 comprise the major findings. They are two ${}_sE_R$ learning curves derived from two different objective measures of performance in the same learning situa-

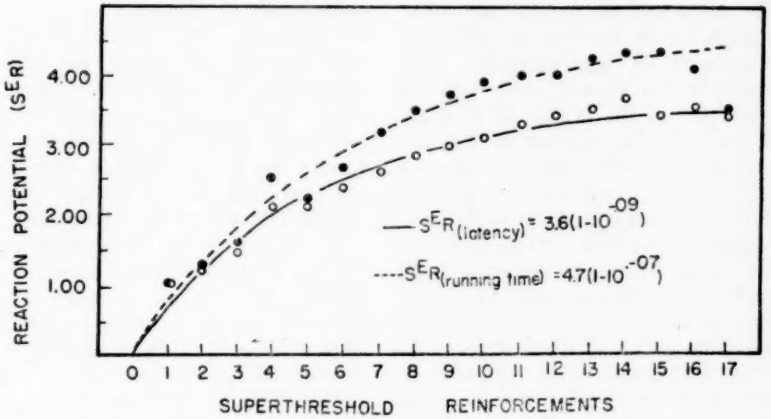

FIG. 3. Two theoretical learning curves derived from the application of the ${}_sE_R$ quantification procedure to two empirical measures of performance in the same learning situation. The two functions are significantly different.

tion and from the same animals. That the two curves are different is obvious by simple inspection. No statistics are needed to establish the true difference between the curves other than this statement: If there is no difference between the curves, the changes of the entire 17 points of the running-time curve falling as they do above the corresponding points of the latency curve are equal to the chances of tossing a coin 17 heads in succession.

More precisely how the two curves differ may be determined by the equations fitted to them. The problem of equation fitting here was somewhat simplified. Since the significance of this study derived from its inclusion within a systematic framework, the choice of an equation type may be such as to facilitate comparison with other parts of the system. Hull (5) has used exclusively the exponential type function to represent habit growth and has related tentatively the constants of this type of equation to certain of the experimental variables. For these reasons the best-fitting exponential type function was found for each set of data. The two equations read:

$${}_sE_R \text{ (latency)} = 3.6\ (1 - 10^{-.09\,N}),$$

$${}_sE_R \text{ (running-time)} = 4.7\ (1 - 10^{-.07\,N}).$$

No greater precision in the approximation of equation constants than is indicated is required to give fair representation to the data and at the same time mediate the quantitative comparison which is the goal of the paper.

To be compared with the above two ${}_sE_R$ equations is the equation found by H.F.G. & Y. (3):

$${}_sE_R \text{ (latency)} = 3.084\ (1 - 10^{-.0205\,N}) + .04.$$

Their equation was found to hold over the course of 60 learning trials. If this equation were graphed for the first 17 trials and plotted in Fig. 3, the resultant curve would fall far below the running-time and latency curves already there.

The relationships among these three functions require analysis and further discussion.

DISCUSSION

1. The large difference between the ${}_sE_R$ (latency) function from the H.F.G.

& Y. study and the present one would be predicted on the basis of difference in amounts of reinforcement.[5] The amount in the present study was .60 gram; that of the H.F.G. & Y. study .08 gm. The asymptotes of the two functions are only slightly different (respectively, 3.6 against 3.1); a larger relative difference appears in the rate-of-approach constant (.09 against .03). The latter difference expresses itself graphically in the fact that the latency curve in Fig. 3 closely approaches its asymptote within 17 trials, whereas the H.F.G. & Y. latency curve requires about three times this number of trials to approach its final limit. If this analysis is true, it leads to the conclusion that the parameter of amount of reinforcement affects the rate-of-approach constant part of the habit strength equation rather than the asymptote part as suggested by Hull (5). This is to be contrasted with a previous finding (14) that when exponential equations are fitted to a family of latency learning curves generated by the amount parameter, it is the asymptote that varies and not the rate-of-approach constant. It is possible that both of these findings are true, since there is no simple one-to-one relationship between an ${}_sE_R$ learning curve and an empirical learning curve both derived from the same data. Comparison of Figs. 1 and 2 shows this very well.

2. If latency and running time are accepted as objective indices of the very same habit strength acquisition, then the critical test of the ${}_sE_R$ quantification methodology as suggested by H.F.G. & Y. (6) fails. This follows from the fact that the ${}_sE_R$ curve for latency differs from the ${}_sE_R$ curve for running-time. Both equation constants for the two curves differ but the heaviest difference lies in the constant controlling the final limit (3.6 for latency asymptote against 4.7 for *RT* asymptote).

3. The test of the methodology need not fail if the measures themselves are questionable. Three aspects of the measures may be questioned: (1) their objectivity, (2) their acceptability as measures of reaction potential, (3) their acceptability as measures of *the same habit strength*. The measures are obviously objective. They are both acceptable as measures of reaction potential on the grounds of convention—they both have been widely used; on the grounds of evidence—a wide range of behavioral processes has been demonstrated by their use; and on *a priori* grounds of generality—the more data that can be incorporated into a system, the better. But the third aspect may be seriously questioned. In fact, the runway habit has been analyzed more than once (1,12) into a chain of at least two components, or sub-habits. An alternative explanation presents itself, then, in place of that of failure of the methodology in question. Latency and running-time may be regarded as two different indices of two different habits within a chain of habits leading to a single terminal reinforcement. The first habit in the chain, coming out of the box, is measured by latency; the second member of the chain, traversal of the runway, is measured by running time.

4. The charge must be faced that this type of reasoning is circular: If the two measures give the same ${}_sE_R$ function, they are two measures of a single habit; if not, they are two different habits. In this way the quantification methodology would always be saved.

[5] Not only the amount of food but the kind of food was different. Observations made by the author on a number of animals revealed that rats receiving cheese as reward in an experimental situation and dog-chow pellets as home-cage food will choose each type of food with about equal frequency when both are available in a neutral situation. This suggests that the difference between the two curves is attributable to the large difference in amount of reinforcement.

The *ad hoc* nature of this reasoning is more apparent than real, because there are differential consequences of there being one or two habits involved. If the ${}_sE_R$ quantification of two measures yields different curves, and it is then asserted that there must be two different habits, then there should be observable differences between these two habits which can be tested. The two habits in the present case, for example, differ in their temporal and spatial distance from the goal. The consequences of one of two habits being farther away in time from a goal are known. The habit which has its reinforcement delayed the longer will be the weaker. From this it can be simply deduced that the latency ${}_sE_R$ learning curve, representing the growth of the coming-out-of-the-box habit, will be exceeded by the running-time ${}_sE_R$ curve, since the latter represents a habit stronger by virtue of its closeness to the goal. This deduction is true. Figure 3 shows the greater strength of the runway habit at every stage of learning. This is shown by the ${}_sE_R$ curves only—not the empirical learning curves of Fig. 1.

There will usually be many other differential consequences of the assumption of two habits. One may have greater inhibitional components and be expected to extinguish faster; or the two habits may have different gradients of stimulus and response generalization; or the serial position of the habits may make a difference. It would require further experimentation with the runway to determine both the presence of these factors and their consequences. However, the difference of delay of reinforcement is a necessary consequence of the physical construction of the apparatus, and the general effect is already well established by the many delays of reinforcement studies (5, p. 163). Hull's theory accounts for more than the direction of superiority of the short-delay member of the chain. A rough quantitative prediction is permitted by the theoretical assumption of a direct relationship between the delay of reward and the asymptote of learning. Inspection of the two ${}_sE_R$ curves in Fig. 3 shows that it is predominantly the asymptote constant which distinguishes the two functions. These findings add support to the analysis of the runway sequence into two functionally discrete habits. They also bear upon the goodness of the reaction potential quantification procedure as a valid method of providing comparable measures of the two habits. Following quantification there appeared theoretically expected differences between the habits which could not be tested with the non-comparable, empirical measures.

SUMMARY AND CONCLUSIONS

1. The reaction potential (${}_sE_R$) quantification procedure set forth by Hull, Felsinger, Gladstone, and Yamaguchi (6) has been applied to two measures of performance of the same animals in the same learning situation to provide a test of the explicit assumption that such a procedure would yield a single function from the two empirical learning curves despite any difference in observed form and magnitude.

2. The empirical learning curves of the two measures—latency and running time—obtained from the Graham-Gagné type of runway were shown to be different. A mathematical statement of this difference was given by the equations found to closely represent the two learning curves. A power function fitted the latency data while an exponential equation described the course of learning as measured by running-time.

3. A number of the statistical assumptions underlying the quantification procedure were tested. They were considered sufficiently tenable to warrant

comparison of the two ${}_sE_R$ learning curves generated by the procedure.

4. The two ${}_sE_R$ learning curves were different. Exponential equations fitted to the two theoretical learning curves showed that they differed predominantly in respect to the constant controlling the asymptote of learning.

The latency ${}_sE_R$ curves also differed from that obtained with a different apparatus by Hull, Felsinger, Gladstone and Yamaguchi in being a much faster rising curve. Hull's theory can be easily adjusted to account for this difference on the basis of a difference in amount of reinforcement.

Such an explanation could not be made of the difference between the ${}_sE_R$ curves for latency and running-time.

5. The methodology under test cannot be said to have failed unless a possible alternative explanation of the results is ruled out.

6. The alternative explanation states that the methodology is valid, but that the two measures are not of the same habit but of two different habits in a chain leading to a single terminal reinforcement. A number of possible tests of this alternative explanation are stated. They involve the testing of the differential consequences of there being two rather than one habit. One such test was possible with the data on hand, and the test confirmed the likelihood of the alternative explanation by predicting the observed difference between the two ${}_sE_R$ curves on the basis of delay of reinforcement.

7. While it is probably true, therefore, that the data on hand do not provide a *critical* test of the reaction potential quantification procedure, they do provide relevant confirming evidence of the utility of the procedure in bringing the empirical facts into closer relationship with a theory of behavior.

REFERENCES

1. BERRY, R. N., VERPLANCK, W. S., & GRAHAM, C. H. The reversal of discrimination in a simple running habit. *J. exp. Psychol.*, 1943, **32**, 325–334.
2. FELSINGER, J. M., GLADSTONE, A. I., YAMAGUCHI, H. G., & HULL, C. L. Reaction latency (str) as a function of the number of reinforcements (N). *J. exp. Psychol.*, 1947, **37**, 214–228.
3. GLADSTONE, A. I., YAMAGUCHI, H. G., HULL, C. L., & FELSINGER, J. M. Some functional relationships of reaction potential (${}_sE_R$) and related phenomena. *J. exp. Psychol.*, 1947, **37**, 510–526.
4. GRAHAM, C. H., & GAGNÉ, R. M. The acquisition, extinction, and spontaneous recovery of a conditioned operant response. *J. exp. Psychol.*, 1940, **26**, 251–280.
5. HULL, C. L. *Principles of behavior.* New York: D. Appleton-Century Co., Inc., 1943.
6. ——, FELSINGER, J. M., GLADSTONE, A. I., & YAMAGUCHI, H. G. A proposed quantification of habit strength. PSYCHOL. REV., 1947, **54**, 237–254.
7. MOTE, F. A., JR., & FINGER, F. W. Exploratory drive and secondary reinforcement in the acquisition and extinction of a simple running response. *J. exp. Psychol.*, 1942, **31**, 57–68.
8. SCHLOSBERG, H., & SOLOMON, R. L. Latency of response in a choice discrimination. *J. exp. Psychol.*, 1943, **33**, 22–39.
9. THURSTONE, L. L. A law of comparative judgment. PSYCHOL. REV., 1927, **34**, 273–286.
10. ——. Psychophysical analysis. *Amer. J. Psychol.*, 1927, **38**, 368–389.
11. ——. Stimulus dispersion in the method of constant stimuli. *J. exp. Psychol.*, 1932, **15**, 284–289.
12. VERPLANCK, W. S. The development of discrimination in a simple locomotor habit. *J. exp. Psychol.*, 1942, **31**, 441–464.
13. YAMAGUCHI, H. G., HULL, C. L., FELSINGER, J. M., & GLADSTONE, A. I. Characteristics of dispersions based on the pooled momentary reaction potentials (${}_s\hat{E}_R$) of a group. PSYCHOL. REV., 1948, **55**, 216–238.
14. ZEAMAN, D. Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 1949, **39**, 466–483.

[MS. received May 16, 1949]

THE APPLICABILITY OF MOTIVATIONAL CRITERIA TO EMOTIONS

BY R. H. WATERS AND D. F. BLACKWOOD

The College of Wooster

In a recent paper Leeper (15) suggests that, contrary to outside opinion, present-day psychology needs not more facts and less theory, but theory tailored to conform more closely to the facts which have already been accumulated. As an example of the sort of theoretical revision which is needed to keep psychology abreast of its growing body of factual material, Leeper examines the treatment given the emotions in some of our widely used texts and argues that their treatment of the emotions as disorganized phenomena is obsolescent. He shows that the terms "disorganized" and "disorganizing" are frequently used in a contradictory manner and that a careful survey of the factual material at our disposal shows them to be poor generalizations. According to Leeper the term "disorganized" has been incorrectly applied to data which actually show reorganization, organization along new lines. Because they arouse and organize behavior along given lines, Leeper holds that the emotions operate as motives and should properly be included within the concept of motivation.

Two other papers have since appeared bearing directly on his proposal. Duffy (11) applauds his stand and discusses some of the theoretical issues involved, such as the difference between emotional and non-emotional behavior and the problem of the nature of emotional processes which would set them off from other motives. Webb (22) also adds other objections to the current treatment of emotion and suggests a possible theoretical orientation for the new view. That neither of these, though contributing significantly to the problem, attack the question along the lines of the present paper will become clear as we proceed.

The suggestion that emotions be included in the category of motivation demands an examination of the concept of motivation to determine whether such an inclusion can be justified. The present paper is an attempt at such an examination. It will be restricted to the applicability of motivational criteria to emotional phenomena.

It should be kept in mind from the outset that a concept is a symbol standing in lieu of a group of specific referents which share some common attribute. The concept "dog," for instance, refers to a group of animals which have some things in common, but all of which are by no means identical. Thus in order to justify the inclusion of the emotions within the category of motivation, it is not necessary to demonstrate that the emotion of fear and the hunger drive are the same thing, but merely to show that fear is similar to hunger in the same way in which hunger is similar to sex need, that is to say, to show that they have comparable dimensions.

Our first move must be to ferret out the criteria of motivation, to determine what common attributes hunger, thirst, sex need and the like have which cause us to think of them as belonging in the same category. These criteria may then be tried on the emotions and tested for goodness of fit. A study of the literature reveals that there have been, in

general, two sets of criteria commonly applied to motives. According to Young (24) these may be designated the behavioral and the physiological.

Motivation is a common concept in psychology today and certainly an important one. Why, then, should there be a difference of opinion as to its nature? The answer to this question is that motivation occupies the position of an intervening variable (12). It is not itself directly observable but is inferred to exist between two units of observable behavior in order to explain a relationship which seems to exist between them. The difference in interpretation depends upon the size of the unit which is observed. For example, according to the behavioral criterion, hunger may be inferred when (a) behavior is aroused, and (b) it is organized in such a way as to lead to the ingestion of food. Many psychologists who insist on a finer analysis of the behavior sequence will not admit the adequacy of this criterion. They will point out that when the rat is observed to enter the food box we are not justified in assuming that he is operating under hunger motivation unless we also know something about his physiological condition. Thus, according to the physiological criterion, it is necessary to demonstrate (a) an atypical tissue condition, followed by (b) gross behavior oriented toward some goal, followed by (c) a change in the tissue condition.

Some writers admit that the physiological criterion is adequate for what are called the biological drives but contend that it is inadequate for the demonstration of higher types of motivation.[1] What, they ask, is the tissue condition characteristic of a desire for social approval? We have, then, a sort of double concept of motivation embracing simple biological drives which have a clear physiological basis and the so-called higher motives in which no such condition is objectively demonstrable.

If a physiological approach is to be used in the area of higher motivation, then the following alternatives present themselves. Some may infer the presence of the physiological condition and make no distinction between drive and motive. They may argue that the physiological basis of higher motives *would* be demonstrated were our techniques sufficiently refined. Others, preferring a broader and safer course, may conceive the higher motives as growing out of biological drives through learning. There is even further division among the latter with some, like Dockeray (8),[2] considering the derivation to be direct and the learning processes involved rather simple. Others, like Allport (1), hold that the genesis is devious and the learning processes so complex that the resultant motives become functionally autonomous, completely distinct from and independent of their roots in tissue need. The two foregoing examples have been utilized as illustrations of extremes along a continuum of learning imposed upon drive. Not all positions have been so extreme, nor have they been completely consistent.[3] Having determined the criteria of motivation and indicated the alternative positions with regard to the relation of drive to motive, we may now

[1] Those who hold strictly to the behavioral criterion will not admit the adequacy of the physiological criterion even in the area of the biological drives. It is beyond the scope of this paper to examine the validity of the physiological criterion. It is sufficient for the purpose of the present analysis to note that it is a widely held position.

[2] Dockeray's position does not actually represent the theoretical extreme of the continuum, but approximates it about as well as any.

[3] Compare Dashiell's treatment of motivation in his 1928 edition (5) with that in his 1937 edition (7).

proceed to an examination of the concept of emotion.

Leeper's protest is not the first statement of dissatisfaction with the concept of emotion as an independent category. Meyer (16), writing in 1933, drew an analogy between the concept of emotion and that of volition, and was so bold as to predict that the term would be cast out of our psychological vocabulary by 1950. (His time is growing short.) Shortly thereafter, in 1934, Duffy (9) suggested that "emotion" simply does not exist as a type of behavior clearly discriminable from other types; and again, in 1941, seven years prior to the publication of Leeper's paper, she (10) argued that in their capacity to release energy in a given direction those phenomena which we call emotions are, indeed, a form of motivation.

In terms of the behavioral criterion it is not difficult to reconcile emotional phenomena with the concept of motivation, nor is it inconsistent with some lines of traditional thought in psychology. Such an approach is generally in harmony with the thinking of the functional psychologists, and lends itself rather well to the evolutionary emphasis for which they have been noted. From time to time a number of writers have considered the motivational aspects of emotion according to the behavioral criterion. Such a concept has been outlined by MacDougall (18) who considered that emotional behavior represented innate response patterns calculated to insure the survival of the organism. A similar approach is inherent in Cannon's (3) early "emergency theory," and in the several approaches based upon the biological concept of homeostasis. These last are somewhat closer to the physiological criterion, but lean heavily upon survival in the evolutionary sense.

The behavioral criterion does not depend upon the essential nature of the phenomena, but rather describes the way in which they operate. For example, we observe that an animal directs its activity toward the ingestion of food and infer that it is motivated by hunger. Similarly, we observe the animal retreats from a dangerous situation and infer that it is motivated by fear, that is to say, that the emotion of fear has served to organize behavior in such a way as to remove or alter the exciting stimulus. With his emphasis on the fact that the behavior aroused by the emotions is organized and purposive, Leeper has aptly demonstrated that emotions may legitimately be thought of as motivation in the behavioral sense, which defines motivation as that which serves to arouse, direct, and sustain activity in conjunction with some goal.

Such broad approaches, it may be argued, lay themselves open to criticism because of their teleological leanings. They cannot be acceptable to those who insist upon a strict physiological interpretation of motivation. For these psychologists motivation must be a matter of here-and-now reduction of an immediate tissue need. Can it be shown that emotions meet the physiological criterion of motivation, that is to say, that they may be thought of as growing out of physiological conditions and that these conditions are subsequently altered as a result of behavior?

Although emotions as motivation do not appear to fit so readily the physiological criterion, there have, nonetheless, been some suggestions made in that direction. This sort of inclination is to be noted in the writings of Woodworth, who says:

"Anyone will unhesitatingly classify as emotions: anger, fear, disgust, joy and sorrow; and as states of the organism: hunger, thirst, nausea, fatigue, drowsiness, intoxication. Now that physiology has revealed

a peculiar organic state in fear and anger, why do we continue to call them emotions and deny that name to fatigue or drowsiness?" (23, p. 234.)

Woodworth has thus indicated that some emotions are comparable, physiologically, to other states commonly called drives. Cannon (4), writing in 1922, catalogued the tissue changes involved in pain, hunger, fear and rage. The inclusion of hunger among the conditions we call emotions appears particularly significant, for is not hunger one of our more commonly referred to biological drives?

In terms of Dockeray s treatment of motivation derived from drive, an approach fairly common in our texts, the emotions can be fitted within such a framework. Dockeray views a drive as a tissue condition which stimulates the organism to random activity until, through trial and error, it hits upon some response which serves to alter the tissue need. This approach is wholly consistent with the usual genetic description of the development of the hunger motive according to which the organism receives sensory impressions from the stomach, is stimulated to a wide variety of activity which brings it into contact with food, the exciting stimulus (tissue condition) is modified, and the organism, if it be a human one, learns to give the sensory impressions from the stomach the name "hunger." [4] When, through learning conceived as being of a rather simple sort (conditioning), the organism comes to associate the specific tissue need with a specific response which will serve to reduce it, it is called a motive. Motive, then, is equal to biological drive plus relatively simple learning. There is some experimental evidence to indicate that emotions as motives can be similarly interpreted. Miller (17), working with rats, used an induced fear as a presumably motivating condition in a learning situation and found that it did indeed operate as a motive in that it (a) produced trial and error activity, and (b) served to reinforce the correct response.

Obviously, further research of a similar nature is indicated. While the outlook, on the basis of this one experiment, is promising, it may be well to look to some of the theoretical implications of such an inclusion. Considering emotions as motivation within this framework may afford us some insight into the existential nature of the emotions themselves. Recall that for Dockeray the drive which directly produces the motive is a tissue condition. The motive then becomes, in a sense, the organism's interpretation of a specific tissue need. Fitting emotions to such a framework is tantamount to an acceptance of the James-Lange theory of emotion (13, 14), which holds that an emotion is the perception of certain bodily changes, or, in more behavioristic terminology, an emotion is the name which we learn to give to a particular pattern of sensory impressions. It need hardly be pointed out that this theory is not currently in good repute since it implies that patterns of internal stimulation in the several emotions must be discriminable one from another and since no such discriminable differences have been satisfactorily isolated. If there are no discriminable differences in patterns of stimulation afforded by the physiological components of the various emotions, then we have a situation somewhat analogous to that found in the field of higher motiva'on. Clearly, if Leeper's idea of emoti s as motiva-

[4] It will be seen that this represents an oversimplification of the mechanisms involved in the hunger drive whose physiological components are known to include much more than stomach contractions alone. Stomach contractions are here used as an illustration of the sort of tissue condition which Dockeray considers to furnish the drive.

tion is to be adopted in conjunction with a concept of motives directly derived from tissue need, then further research is needed in the physiology of the emotions, calculated to demonstrate discrete patterns. This is an area which has already received considerable attention and, as has been noted, the results, generally, have been negative.

Fortunately there is another alternative which, while it compromises to some extent the strict physiological viewpoint, effectively circumvents the difficulties inherent in Dockeray's approach. This alternative invokes Allport's idea of functional autonomy. Here motives may also derive from biological drives with a basis in tissue need, but the learning processes involved are so complex that the resultant motivation is completely independent of its physiological roots. The emphasis in functional autonomy is upon the divorcement of the motive from the tissue need involved in its genesis. Thus, when a motive has, through complex learning, become autonomous, it is to all practical intents and purposes psychogenic and one is no longer obliged, when dealing with higher order motives and those areas of emotion which are hazily defined, to demonstrate the present operation of an atypical tissue condition. Hence, as Leeper has held, the theory of emotion which says that emotions are, indeed, psychogenic, and which Munn (19) considers too naïve for serious treatment, may in the final analysis be the least naïve of all the theories which have been advanced.

It may be argued that in invoking functional autonomy to circumvent the need for the establishment of specific tissue conditions in all cases, we have left the strict physiological criterion and have arrived by indirection at a behavioral approach. The only answer to this must be that any theoretical area in psychology must be broad enough to cover all of the involved phenomena. As Hull (12) has said, when theory and fact conflict, theory must always give way.

From the foregoing review of the status of emotion and motivation, it appears that the concept of emotions as motivation can be conditionally reconciled with a physiological approach to motivation, the conditions being these: In order to insure consistency those who hold to the physiological criterion of motivation will be faced with the alternative of either (a) instituting more intensive research with an eye to the establishment of definite patterns of atypical tissue conditions both in the so-called emotions and in the area of higher motivation, or (b) expanding their theoretical framework to include the concept of functional autonomy. Acceptance of the Leeper hypothesis will serve to combine two vague areas of investigation into one and offer positive suggestions for research in this one area which may lead not only to clarification of the phenomena in question, but also to a narrowing of the gap between those who approach the problem from different directions.

The lines of research indicated by such a concept embrace not only investigation of the physiological components of motives and emotions but also the operational aspects of both. We speak, for instance, of levels of motivation which are presumed to vary directly with goal deprivation. This raises the question of the possible quantification of emotions along similar lines. Again it has often been noted that extreme emotion appears to operate in contradiction to a motivational interpretation. At the same time there has been some suggestion that a similar situation exists in motivation (20). Too high a level of motivation may inhibit goal-directed activity.

Further consequences of the adop-

tion of Leeper's hypothesis will probably involve such problems as the reconciliation of the genetic development of emotions according to Bridges (2) with some scheme of the genetic development of motives, a treatment of Dashiell's (6) suggestion that emotions may be nothing more than socially determined interpretations of environmental conditions, and investigation in the area of the acquisition of new motives and emotions.

Of course, the final word on the inclusion of emotions within the category of motives must rest with the experimental findings in the above indicated areas. In terms of the narrower topic assumed by this paper, however, it appears that the hypothesis may be tentatively accepted.

BIBLIOGRAPHY

1. ALLPORT, G. W. *Personality*. New York: Holt, 1937.
2. BRIDGES, K. M. B. A genetic theory of the emotions. *J. genet. Psychol.*, 1930, **37**, 514–547.
3. CANNON, W. B. The interrelations of emotions as suggested by recent physiological researches. *Amer. J. Psychol.*, 1914, **25**, 256–282.
4. ——. *Bodily changes in pain, hunger, fear and rage*. New York: D. Appleton Co., 1922.
5. DASHIELL, J. F. *Fundamentals of objective psychology*. Boston: Houghton Mifflin, 1928.
6. ——. Are there any native emotions? PSYCHOL. REV., 1928, **35**, 319–327.
7. ——. *Fundamentals of general psychology*. Boston: Houghton Mifflin, 1937.
8. DOCKERAY, F. C. *General psychology* (Rev. ed.). New York: Prentice Hall, 1936.
9. DUFFY, E. Emotion: an example of the need for reorientation in psychology. PSYCHOL. REV., 1934, **41**, 184–198.
10. ——. An explanation of "emotional" phenomena without the use of the concept "emotion." *J. gen. Psychol.*, 1941, **25**, 283–293.
11. ——. Leeper's "motivational theory of emotion." PSYCHOL. REV., 1948, **55**, 324–328.
12. HULL, C. L. *Principles of behavior*. New York: D. Appleton Century, 1943.
13. JAMES, W. What is an emotion? *Mind*, 1884, **9**, 188–205.
14. LANGE, C. G. The emotions. In *Psychology Classics* (K. Dunlap, Ed.). Baltimore: Williams and Wilkins, 1922.
15. LEEPER, R. W. A motivational theory of emotion to replace "Emotion as a disorganized response." PSYCHOL. REV., 1948, **55**, 5–21.
16. MEYER, M. F. That whale among fishes —the theory of emotions. PSYCHOL. REV., 1933, **40**, 292–300.
17. MILLER, N. E. Studies of fear as an acquirable drive; I. Fear as motivation and fear reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, **38**, 89–100.
18. MACDOUGALL, W. *Social psychology*. Boston: Luce, 1921.
19. MUNN, N. L. *Psychology*. Boston: Houghton Mifflin, 1946.
20. TOLMAN, E. C., HALL, C. S., & BRETNALL, E. P. A disproof of the law of effect and a substitution of the laws of emphasis, motivation and disruption. *J. exp. Psychol.*, 1932, **15**, 601–614.
21. WATION, J. B. A schematic outline of emotions. PSYCHOL. REV., 1919, **26**, 165–196.
22. WEBB, W. B. A motivational theory of emotions. . . ." PSYCHOL. REV., 1948, **55**, 329–335.
23. WOODWORTH, R. S. *Experimental psychology*. New York: Holt, 1938.
24. YOUNG, P. T. Motivation. In *The Encyclopedia of Psychology*. (P. L. Harriman, Ed.) New York: Philosophical Library, 1946, 384–393.

[MS. received May 27, 1949]

THE NATURE AND FUNCTIONING OF WANTS [1]

BY EDWARD C. TOLMAN

University of California

When we examine the lists of motives, or wants, cited either by common sense or by the system-makers in psychology, we find items which seem to fall into either one or the other of two extreme categories. On the one hand, we find lists of the so-called "basic drives" or "needs," such as: hunger, thirst, sex, aggression, need for affiliation, need for dominance, and the rest. Or, on the other hand, we find statements about very specific and immediate wants such as: the want for a piece of candy when a box or dish appears, the want for a cigarette upon seeing a package handy, the want to get out of the range of a given too hot fire, etc., etc., etc.

Furthermore, corresponding to the emphasis on the one or the other of these two kinds of item, we find two theoretical points of view. There is the type of theory which stresses the importance of the basic drives or needs as the only fundamental, energy or tension providing sources in the organism. This type of theory seeks to *derive* all the immediate, more situationally described, wants from these basic drives or needs. The theorists whose names suggest themselves in this connection are, of course, Freud (4), McDougall (16), Murray (19), Sanford (23), Maslow (18), Goldstein (5), and also Hull (9). And secondly, there is the kind of theory (the names which suggest themselves are Koffka [11], Lewin [14], Krech [13], Crutchfield [13], MacLeod [15]) which tends to take the immediate situational want as its paradigm and conceives the functioning of the basic drives and needs as merely highly charged instances of immediate situational wants. This latter type of theory stresses the immediate environmental context. And, whether the want in question be that of getting to food or to a sex-object or merely that to hang a crooked picture straight or to pick a pin up off the floor, the important determiners as seen by this second sort of theory, which I shall call the "neo-Gestaltist" theory, lie in the "dynamics" of the immediate "phenomenological" field. The basic laws of motivation are believed by this second group to be analogous to the Gestalt laws of perception—such laws, say, as: "Closure," "Prägnanz," "Better Figure" or the like.

The point of view to be defended in this paper is a compromise one. I would hold with the neo-Gestaltists that the *functioning* of a want, *as an immediate event, is* always an activity going on in what may be called an immediate "phenomenological" or, as I would prefer to conceive it, "behavior" field—in, that is, the present situation as immediately registered by the organism as a set of behavior possibilities. (See Fig. 1.) And this I would hold to be true whether the given instance be a case of hunger, or of sex, or of the mere opening of a door or of the mere removing of a speck from the floor. The immediate process in each instance would be some sort of a "closure"-like, or "better figure"-like phenomenon taking place in the immediately registered

[1] Presented to the graduate students of the Universities of Iowa, Chicago, and Michigan and of Western Reserve University.

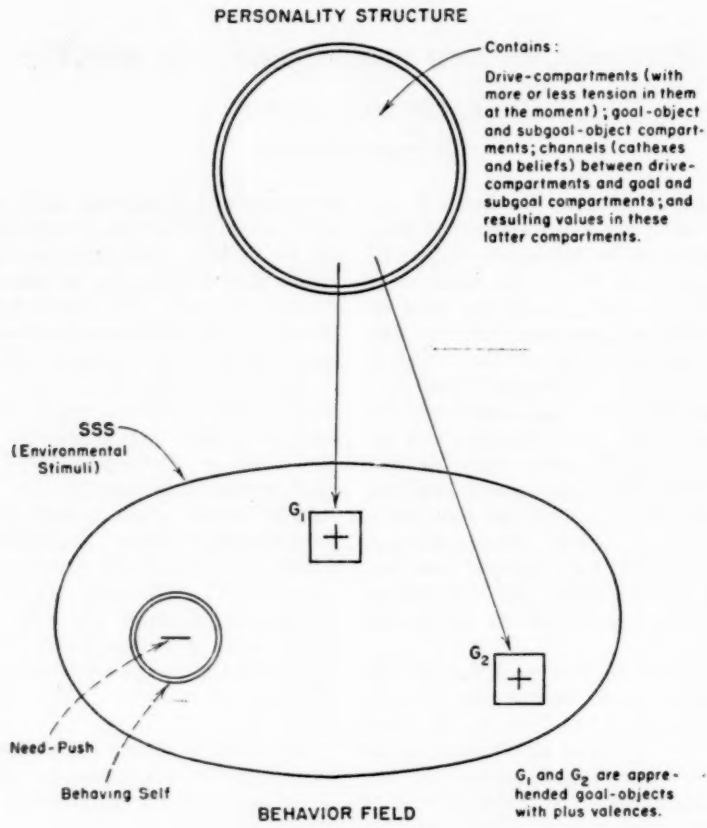
PERSONALITY STRUCTURE
Contains:
Drive-compartments (with more or less tension in them at the moment); goal-object and subgoal-object compartments; channels (cathexes and beliefs) between drive-compartments and goal and subgoal compartments; and resulting values in these latter compartments.
SSS
(Environmental Stimuli)
G1
G2
Need-Push
Behaving Self
G1 and G2 are apprehended goal-objects with plus valences.
BEHAVIOR FIELD

FIG. 1

—i.e., apprehended—behavior field.[2] A valenced goal is registered, apprehended, or valenced goals are apprehended, as located there or there in the field. And the organism apprehends also what I shall call a "need-push" in his "behaving self"—this behaving self with its need-push is likewise apprehended and apprehended as located at its own different place in the behavior field. The apprehended need-push pushes the apprehended behaving self towards the apprehended valenced goal or goals. The result is a bringing of the behaving self with its need-push and the valenced goal or goals nearer together either by the organism's actually locomoting towards the goal or (if locomotion proves impossible or not permitted or if the motivation be irrelevantly en-

[2] From here on I shall, because I think it is more euphonious, substitute "apprehended" for "registered." But by "apprehended" I do not mean a necessarily conscious process any more than I would by the term "registered." "Apprehended" will mean, as I use it, something grasped by or registered on the nervous system and from my point of view such a registering or apprehension may be, and I suspect quite often is, wholly unaccompanied by consciousness.

hanced) by some sort of a cognitive distortion of the apprehended distance between the behaving self and the goal.[3]

However, while I would thus agree that the functioning of any want, whether basic or superficial, always obeys some final "dynamic" field laws, I would also hold that for the complete understanding of why a want appears and functions in a given case there have to be considered also the peculiar physiological, glandular, nutritional and past training effects acting upon the particular individual at the particular time. And it is when one turns to these individual, often historical, determiners that one finds it useful to conceive not only of immediately apprehended behavior fields but also of more enduring factors in the organism which I would summarize under the concept of a personality structure. Such a personality structure would be a complex of underlying "drives" or "needs" plus attachments of these drives or needs (as a result of past experience) to goals and subgoals. (See Fig. 1.) In short, one usually has to consider a relatively enduring personality structure with its inlying drives, drive-attachments and resultant values to explain why in any given instance the need-pushes and the valences in the given behavior-field are what they are. Individual constitution, individual history and present-acting metabolic or glandular conditions seem to be especially important for determining the strengths of present drives, the strengths of the attachments of such drives, and the final resultant values.

I would assume, that is, some list of innate *drives* or *needs*, although I do not know *now* just what this list will finally turn out to be. Each basic drive or need I would suppose to have different average strengths for different individuals due to differences of innate constitution[4] and due also, perhaps, to differences of early training. I would assume further that in any one individual a given drive may suffer (some types of drives more than others) periodic enhancements and diminutions due to varying nutritional and glandular conditions and disturbances. Such temporal variations in the strengths of the drives are particularly obvious in the case of what have been called by Murray (19) the "viscerogenic needs"—*i.e.*, hunger, thirst, sex, and the like. Such viscerogenic needs or viscerogenic hungers exhibit cyclical rhythms dependent upon physiological processes and upon the interactions of environmentally controlled maintenance-schedules with these physiological processes. For example, we in America, as a result of physiology plus culturally imposed maintenance-schedules, normally tend to become food-hungry three times a day. And what holds for food-hunger tends to hold in somewhat analogous fashion for the other viscerogenic hungers, such as: thirst, sex, maternal drive (in lactating mothers), toilet-seeking, and the rest.

My next point would be that the drives are attached, "channeled," or as I shall call it from now on "cathected," to specific types of final goal-object either innately or as a result of experience. Further, I shall hold that additional channelings develop between such types of potential final goal-objects and still other types of objects as possible means or subgoals. These latter channelings I shall term "beliefs" (mostly instrumental in character) and I shall

[3] In this connection we may recall the C[illegible]hfield experiments (1) in which rats after being trained to run a certain distance under one degree of hunger overshot the mark when their hunger was increased.

[4] See Calvin S. Hall's experiments on breeding strains of rats with different average strengths of timidity and different average strengths of aggressiveness (6, 7).

maintain that it is the cathexes and beliefs constituting relatively simple or relatively long and complex chains of connections to the underlying drives which give rise finally (together with the presented stimuli plus the laws of perception) to the apprehensions of need-pushes, of valences, and of directions and distances in any actual behavior field.

Return, for the moment, to the situational wants, which are emphasized by our neo-Gestaltist friends—such wants, say, to list a few more, as those to "sharpen a pencil," to "unravel a piece of string," to "sing a specific tune," to "select and buy a new shirt," and the like. All such wants are conceived by me as having to be traced back in the last analysis through cathexes and instrumental or other types of belief to underlying drives or needs. However, I hasten to add that such a tracing back will often be relatively difficult or even (for lack of complete historical knowledge concerning the individual) to all intents and purposes impossible. Furthermore, such a tracing back, even if achieved, may well yield but little useful information. For, in most everyday instances, practically all the individuals in a given culture will have very similar maintenance-schedules, will have been exposed to the same learning processes and hence will have all arrived at practically the same sets of cathexes and instrumental and other beliefs, and will, therefore, respond in most ordinary situations with practically the same resultant need-pushes and valences. Put any middle class college male in a store before an array of shirts and he will behave practically like every other middle class college male. That is, his behavior can, for all practical purposes, be considered as governed wholly by the immediate situation and not by any deep-lying drives and cathexes and instrumental or other beliefs. Given a common culture and typical individuals in that culture and the everyday goals and subgoals of all such individuals will be practically identical. Any attempted examination of the specific drives and cathexes and beliefs underlying one of these individual's choices may well, therefore, seem superfluous and silly. For practically all variations in the behaviors of these individuals will be correlated primarily with differences in the environment. There will be no need of taking into account basic drives and cathexes and beliefs. For these will be the same for all such individuals.

However, put very different sorts of persons into the same situation and a quite different picture appears. Thus, if we introduced into our haberdasher's an unusually rich student, or an unusually poor one, or a visiting chief from a native culture, or a very sexually aroused, showoff male, then each of these might well tend to select very different sorts of shirts. And it would appear at once that in order to explain these differences of preference (*i.e.*, of resultant valence and of need-push) in these different men, it would be necessary to consider not only the objective environmental situation, which would be practically identical for them all, but also the differences in their basic drive strengths at the moment, and in their individual histories which would have produced different cathexes and different beliefs attached to such drives.

But let us now analyze "cathexes" further. The term "cathexis" has, of course, been borrowed from Freud (4). I have, however, somewhat departed from his usage. By a cathexis I shall mean not the resultant loading of the given type of goal-object by the drive (this loading I shall call value, see below) but rather the innate or acquired connection (or channel) between the given drive, say a viscerogenic hunger (such as food-hunger or thirst-hunger,

or sex-hunger) and the given *type* of goal-object. The result of such a cathexis (*i.e.*, such a channel) will be that when the given drive is in force and this cathexis channel is deep, that is, highly permeable, there will be a flow of energy from the drive to the compartment for the given type of goal-object. And this will result in what I call a corresponding *positive value* for that *type* of goal-object. Such a positive "value" will bring it about that when it, this value, is high, the given individual will at that moment be especially likely: (1) to apprehend in his behavior space any concrete instance of such a goal-object provided by the environment; and (2) to endow such an instance with positive "valence." It will be noted that I am using "value" to designate the positive readiness for a *type*, or *class*, of objects, and that I am using "valence" to designate the resultant positive readiness to go to a specific instance of such a class of objects when such a specific instance is apprehended in a particular behavior space. A value is carried around by the organism. A valence is a specific embodiment of this value in an apprehended concrete behavior space.

Consider next "beliefs" and for the present primarily instrumental beliefs. It is my contention that the organism in addition to carrying around in his personality structure attachments of types of goal-object to drives—*i.e.*, cathexes—also carries around attachments of types of subgoal to types of final goal or to types of further subgoal. It is all these attachments, or channels, which I call "instrumental beliefs." These "instrumental beliefs" are, I believe, primarily a product of learning. But further experimentation needs to be done. Examples of some of the probably millions of common, everyday instrumental beliefs which practically every middle-class American has learned and carries about with him might be: "restaurants and/or kitchens lead to food," "schools lead to education," "the paths of speeding automobiles lead to injury," "gas stations lead to toilets," "electric switches lead to lights," and so on and so on and so on.

In addition, however, to (or perhaps better as a sort of encompassing matrix for) many such relatively specific instrumental beliefs there would be still more highly generalized, more abstract and in human beings eventually codified and highly elaborated sets of beliefs or rules which an individual inherits or acquires. These would be general laws, such as those of space, time, quantity, causation, verbal logic, etc. Even animals have some such substratum of such general innate or acquired rules about space, time, quantity, and quality. The animal's general rules are simpler than ours and are not elaborated into mathematical or logical systems. But, nevertheless, a considerable part of what we seem to be interested in doing in our animal studies is in discovering just what general notions, rules, *i.e.*, "instrumental beliefs," as to space, time, quantity, gravity, and qualities our rats and chimpanzees do have or are capable of acquiring.[5]

But consider now, more specifically, the motivational significance of such rules or instrumental beliefs, general or specific. My proposition is that whenever the far-terminal of an instrumental belief is positively valued, the near-terminal of the belief becomes by virtue of this belief (*i.e.*, channel) also positively valued. In other words, instrumental beliefs are similar to cathexes

[5] What I am here calling "general" instrumental beliefs I called in a previous article (27) "field cognition modes." These general beliefs or field cognition modes would also seem to be closely related to what Harlow (8) has recently called learning sets.

in that they are to be conceived as in the nature of channels through which values may be conducted. Whereas a cathexis may be said to be a *final* instrumental belief which operates to bring it about, that if some type of drive-relief is positively valued, a certain type of final goal-object will therefore also be positively valued, an instrumental belief (or it may be that a whole nexus of instrumental beliefs or a belief-set) operates to bring it about that if some such type of final goal-object is positively valued, then such and such types of instrumentally connected subgoals will also be valued.

By this time the reader must have become restive and be asking himself whether I have any possible objective, operational ways of defining these constructs of drives, values, cathexes, beliefs and the other constructs which I have not yet talked much about, of apprehended behavior fields, behavings-selves, valences and need-pushes. I do not pretend that I have any good final definitions. I can, however, suggest the directions in which, I think, such definitions should be searched for.

(1) *Drives* (or *needs*). The operational definition of a drive, or need, is to be found, I believe, in the propensity of an individual to perform a characteristic type of consummatory response. Thus, food hunger would be defined by the propensity for eating, thirst by that for drinking, sex by that for the sex act, dominance by that for performing the act of dominating over other individuals, maternal drive by that for the act of nursing or otherwise caring for the young, curiosity by that for the acts of investigating and examining, and so on.[6] Thus one possible empirical technique for identifying, and measuring the strength of, any such drive or need would consist in putting the individual in a "standardized situation" in the direct presence of a "standard" goal-object appropriate for the given drive and then noting the appearance of the consummatory response itself and measuring it quantitatively by its vigor, duration, frequency, latency, or the like. A second possible technique for measuring the strength of the drive would consist in putting the individual in a situation with a standardized distance, or obstacle, between him and the standard goal-object and measuring some quantitative aspect, again such as frequency, latency, duration, etc., not of his consummatory response as such, but of his approach-to-the-goal response. It is, of course, this latter *approach technique* which is used by the Columbia Obstruction Box, whereas it seems to be rather the strength of the consummatory response itself which was used by Wiesner and Sheard (29) in their counting of the number of retrievings of young to the nest which were carried out by a mother rat; and also by Maslow (17) and Nowlis (21) when they rated dominance in the monkeys and chimpanzees by pairing each animal with every other and observing in each such pair which animal dominated over the other in the situations of getting to food or to the home cage.

Of course many problems and difficulties, involved in such attempts to measure the strengths of drives by strengths of consummatory responses or of approach responses, immediately

[6] This fact that I am identifying drives by the character of consummatory responses, means, I hope, that I am escaping the pitfall of finally ending up with an innumerable number of drives. For there are, I believe, only a relatively limited number of distinctive consummatory responses. There will not be as many drives as there are types of goal-object but only as many drives as there are distinctive different ways of behaving to goal-objects. And I believe we may some day be willing to settle for some number not much more than twenty.

arise. There is the problem of scale-units. Do twenty crossings of the grill in twenty minutes indicate a drive which is twice as strong as one which evokes only ten crossings in twenty minutes? But this problem of the equality of scale-units is one which has plagued all attempts at psychological measurement and is not peculiar to such measurements as these of drives. It is equally a problem for the hard-headed psycho-physicists and the hard-headed mental testers.

A second type of difficulty arises out of whether the conditions introduced into the measuring situation itself—such as, for example, the use of an electric grill in the Columbia Box—may not introduce an interacting and hence falsifying factor. Again, however, this is a question which can be attacked and eventually solved only by thinking up other ways of measuring. It is the question as to whether, or in how far, the nature of one's measuring instruments and procedures themselves enter into and distort the measurements. This is a problem for physics as well as for psychology. (Consider, for example, the Heisenberg Uncertainty Principle.) We can make our final decisions about it only by inventing other types of measuring stick and seeing whether or not they give the same or seemingly better, more meaningful, results and concepts.

Yet again, another difficulty, especially in the case of human beings, is that these latter tiresome animals seem to be able to cover up, repress, inhibit, or distort their drives so that when you put them in the presence of the standard goal-object they refuse to go to or consume it, in any obvious, overtly measurable way—in spite of the fact that *you* know that they do suffer from the given drive or need. Thus, for example, one can put a hungry, but overly well-mannered child in the presence of food, and he may refuse to eat. And it is a commonplace that what can thus be brought about in our culture by social probibitions in the case of hunger can be brought about much more seriously in the cases of sex and aggression. It is, of course, in the attempt to meet such difficulties that the projective techniques have been invented; but as ways of measuring momentary drive strengths these projective techniques are, of course, still far from satisfactory. Nevertheless, they do offer a beginning (see, for example, the experiment of Sanford [24] which developed a projective test for measuring immediate hunger).

(2) Turn now to our next concept—that of *values*. A *positive value* (and in the present paper I am considering positive values only, although a complete discussion would necessarily involve negative ones as well) I would measure operationally by the magnitude of the propensity to get to a specific type of goal object. Some quantitative aspect of this getting-to would constitute my measure of a positive value. But it will be remembered that such a measuring of the magnitude of the getting-to response was one of the techniques just suggested for measuring the strength of the drive. Hence, in the case in which the type of object in question is the *standard* goal-object, the quantitative magnitudes of the strength of the drive and the quantitative magnitudes of the positive values of the goal-object will be one and the same. (The significance of this will be clearer in a moment, I hope.)

Both concepts, that of drive and that of value of goal-object, are needed in order to allow for types of goal which evoke weaker approach responses than does the standard goal. The hungry rat approaches a standard food with a certain degree of vigor. This will operationally measure the strength of his hun-

ger drive. He approaches other foods, however, with either less or more vigor. These latter foods will have smaller or larger positive values given the same strengths of drive. The strength of the hunger drive at any given moment is defined by the vigor with which the individual will approach the standard food. And the vigor with which he approaches any non-standard food will measure his *value* for this other non-standard food.

Consider also the sort of case in which a given type of object serves as goal for more than one drive. The strength of the approach propensities (the value) of such an object may then prove exceptionally strong. For example, a given kind of food may have prestige value in addition to hunger value and hence be gone to more strongly than the standard food. A child's value for ice cream may perhaps be an example of this. Ice cream may be approached both as food and as a socially approved prestige goal.

(3) Consider next operational definitions for *cathexes* and for *instrumental* and *other beliefs*. These are all, as I have said, to be conceived as hypothetical *channelings* between successive value compartments in the personality structure. These channels act so that when one compartment is filled with value any other compartment connected to it by such a channel also tends to become filled.

A. *Cathexes*. A cathexis is a channel between a drive-subsidence compartment and a goal-object compartment. Our immediate task is how to arrive at an operational definition and measurement of such a channel. It is to be recalled that we suggested that we could define and measure the strength of a drive by measuring the getting-to responses relative to the "standard" goal for that drive. Another way to put this is to say that, by definition, I am asserting that the degree of cathexis for a standard object always has the value of "one." For, if we plot different magnitudes of the drive as strengths of getting to a standard goal as our abscissae and plot the corresponding same magnitudes as the positive values of the standard goal as our ordinates, we would necessarily (since both sets of figures will be the same) get a straight line at a 45° angle. And the tangent of this angle would have the value of one. In other words, I shall define operationally the strength of the cathexis to any type of goal by the tangent of the angle made by the line resulting from plotting the magnitudes of the measures for getting-to such a goal against the magnitudes of the measures for getting-to the standard goal. I shall assume, further, that goals other than the standard may be cathected either more or less strongly. Hence the slopes of the lines for those other goals will be greater or less than 45° and the tangents of the angles corresponding to them either greater or less than one. To recapitulate—the standard goal will, by definition, have a cathexis of one and other types of goals will have cathexes either greater or less than one.

B. *Instrumental beliefs*. An instrumental belief I conceive as in the nature of a channel (or possibly a set of channels) between the compartment for a type of goal and that for a type of means or subgoal, or as a channel between two successive subgoals. If the more final one of these goals or subgoals has positive value, the nearer, less final one, will also have positive value conducted to it depending upon the strength of the belief (*i.e.*, the goodness of the channel). How do we operationally discover and measure such beliefs? Theoretically the answer is simple. Practically, however, given the complexity of the human animal and

its covering-up capacities, the uncovering of a man's beliefs and the attempt to measure them is usually a very tricky business, for which at present we have mostly only relatively indirect and fallible procedures, such as interviews and projective techniques. Theoretically, we can proceed after the analogy of cathexes. We can discover that the positive values of object Type X (that is, the propensity to get to this x type of object) go up and down with the positive values of object Type Y. We plot the one set of values against the other and get the slope of the curve. This slope will usually vary from zero to a maximum of 45 degrees, in which latter case the instrumental belief (*i.e.*, "that *x*'s are the way to get to *y*'s") may be said to be 100 per cent, in the sense that the organism believes completely and absolutely that the reaching of an object of Type X will lead him to an object of Type Y.[7]

But any such certain and neatly isolated instrumental beliefs are probably very rare in human beings. Rather, given our shifting experiences and our complicated personality structures, it becomes very difficult to uncover what are the manifolds of instrumental beliefs upon which, for example, your and my respective positive values for the study of psychology depend. Detailed case histories, projective techniques, and depth interviews would all have to be tried, and even then the final plotting out of your and my respective belief systems would still be largely a product of clinical insight and intuition. The situation would be the same as that in which students of prejudice and ethnocentrism have found themselves involved. They have discovered (see, for example, the recent articles by Else Frenkel-Brunswik [?] and Frenkel-Brunswik, Levinson and Sanford [3]) that in seeking to discover why some individuals have strong negative values for minority groups and others have positive values for them, they, the investigators, had to uncover, by all the depth methods at their command, the values of their subjects with respect to a wide range of other entities, especially their plus and minus values with respect to parents and authority figures, in-groups and out-groups, sex-roles, manners, crime, astrology, acts of aggression, etc., etc., etc. And, then, the final spelling out of the instrumental or other belief-connections between the values for all these other entities and the resultant positive value for majority groups seemed to be a matter of good sense and intuition and not one as yet capable of precise experimental confirmation.[8]

C. Finally, at this point, I need to introduce an additional concept, that of equivalence beliefs. (See also [27].) For it would appear that when we try to untangle the nexus of connections by which such a positive value as that for "being a psychologist" or that for "one's own group as contrasted with minority groups" is determined in a given individual, we seem oftentimes to find that the value for one type of entity derives from that for another type of entity not as a result of a true instrumental rela-

[7] The intriguing possibility also suggests itself that fixations, stereotypes, and prejudices may perhaps be considered as beliefs that are more than 100 per cent strong. In such cases the values for x's would be stronger than the values for y's even though they depend upon the latter. Finally, it may be noted that what I am here calling "instrumental beliefs" I formerly called "means-end readinesses" (26).

[8] At this point it should again be emphasized that a complete statement of the doctrine I am presenting would require the inclusion of negative values (for example, the negative values for minority groups). I have had to leave a discussion of negative values out of this account for lack of time. But my system with a little more elaboration can really take care of them.

tionship but rather from the acceptance of an identification, *i.e.*, an equivalence, between the two. The operational evidence for such an equivalence belief would be found (1) in the fact that, as in the case of an instrumental belief, the positive value for the one type of entity would be found to go up and down with that for another type of entity, but also (2) in the fact (not true for an instrumental belief) that when an instance of the equivalent entity is reached, there would be a reduced remaining propensity to go on to an instance of the original entity to which it is equivalent. The equivalent would, in short, be accepted, as in some degree, a *substitute* for that for which it is taken as equivalent, in a way in which a mere subgoal or means is not accepted as a substitute for the final goal.

This last point emphasizes a feature about my whole conceptual system concerning drives and values which has not been made sufficiently explicit. It appears, in short, that when a type of entity has positive value because it is either a type of goal or a type of subgoal, this means not only that there is a propensity to get to any instance of it and to experience an immediate discharge in the fact of getting to such an instance, but also that there remains a propensity to use such an instance to go on to instances of the next subgoal or goal in the instrumental hierachy. Even an instance of a final goal provides not merely a discharge of tension upon being reached, but it is then also manipulated by the consummatory activity as the way to arrive at the still more ultimate goal of drive subsidence. In the case of an equivalence belief, on the other hand, getting to and manipulating the type of entity which is accepted as "equivalent" relieves the need-push for this latter, even though it does *not* get the organism on towards the final goal and does not help in relieving his ultimate underlying drive tension. Yet the organism may persist in getting to the "equivalent" because of his immediate relief, and consistently fail to discover that his underlying drive as such is not relieved. The individual seems to have, in short, as a result of an equivalence belief, what has sometimes been called an insatiable neurotic need.

Finally, it is obvious that, if depth interviews and all kinds of fancy clinical techniques are necessary to unravel an individual's specific instrumental beliefs, such interviews and such fancy techniques will be all the more necessary to unravel the individual's equivalence beliefs.

(4) Finally, let me turn now, very briefly, to the operational definitions of my remaining concepts: the *apprehended behavior field* with its qualities, directions and distances, the apprehended *behaving self*, and apprehended *valences* and *need-pushes*.

My concept of an *apprehended behavior* field is patterned very closely after Lewin's "outer life space" (14). The apprehended behavior field is the form in which any immediate environment is perceived and/or remembered and/or inferred by the organism at a given time. Such a construct of a psychological life space or apprehended behavior field as distinct from the true "geographical environment" (to use Koffka's term [11]) undoubtedly arises in the first place as a result of the alarums and excursions of each individual's own everyday experience. The environment *appears* to each of us to be so and so, but we examine further and discover the facts of illusion, and of our own perceptual, memorial and inferential insufficiencies. We discover the difference between appearance and reality—between our immediately apprehended behavior spaces and the real

environments, which latter we discover from communal investigations by ourselves and our fellows, using meter sticks and all the paraphernalia of both the macroscopic and microscopic physical and social sciences. When we turn to the study of the lower animals, it becomes more than obvious that the apprehended behavior fields of these animals must be in many ways quite different from ours—given one and the same geographical environments. In short, we are reinforced in the distinction between the true geographical environment of any individual X (man or animal) as described by a community of independent observers provided with the measuring tools of the sciences and the apprehended situation (or behavior field) of Individual X as it is indicated by his behavior in some specific segment of time.

But still the question remains as to how we can arrive at operational definitions of such behavior fields and of their contents.

Here it must be admitted we are at present in a pretty unsatisfactory stage. We again find ourselves today relying in large part on empathy, intuition, hunches, introspective phenomenology, and various other subjective, non-operationally definable procedures. The beginnings of a more objective, more operational way for arriving at the probslowly being developed, especially by able behavior fields for given organisms at given times are, however, I believe, animal psychologists.

The probable *perceptual qualities* in a rat's field are being arrived at by discrimination experiments which relate stimuli and patterns of stimuli (physically defined) to discriminable differences in behavior. We learn not only about the possible discriminanda in the rat, but also such facts as that a horizontal row of dots looks like a solid horizontal line (Krech, 12). When all the necessary work has been done, we will eventually be able to postulate what qualities, what discriminanda, arrayed in what *Gestalten* a rat will probably apprehend when presented with a given array of present and past stimuli.

As to *distances* and *directions*, our maze work is just now getting to this point. It is to be hoped that in the near future with the use of sound-proof and light-proof domes (in which we can spot visual and auditory cues at will) we shall be able to discover the basic principles of a rat's perception of spatial distances and directions. We shall be able to discover, that is, what arrays of spatially located cues are going to lead to such and such abilities or non-abilities to take a short-cut, to take an *Umweg*, to take the shorter of two paths, and the like.

As to the rat's or the chimpanzee's perception of its *behaving self*, I confess I am stumped. However, I would guess that the uncovering of the principles of the perception of spatial distances and directions would probably also uncover something to the effect that these perceptions radiate out from some sort of an identified and moving locus, *i.e.*, a behaving self.

Next, as to *need-pushes* and *valences*, The concept of need-pushes as perceived in the behaving self has been forced upon me by a number of considerations (including, I have to admit, my own introspection). A need-push is, so to speak, the activating push in a concrete behavior field toward a valenced goal and it may or may not correspond in amount to what, by other means, we would define as the underlying value resulting from the given drive or need. Thus, for example, there are some very surprising results beginning to appear with rats. It looks from some of the recent latent learning findings—including those of Spence and Lippitt (25), Kendler and Mencher (10) and of

Walker (28), and some things we are now doing at California—that even though a rat's drive be changed from thirst to hunger, his need-push may not, for a while, change correspondingly. The water place may still be valenced, even though the drive as measured by our original definition is now one of hunger. In other words, under some conditions, rats, and men too, do not seem to have need-pushes which correspond to their actual drives. (And also I would remark, parenthetically, they often experience valences which do not correspond to their actual values.) Why this happens, how much it may be a function of early childhood maintenance schedules involving deprivations or non-deprivations, or how much it may result from disrupting conflict situations such, say, as that involved in the use of doors in the original Spence and Lippitt (25) latent learning experiment and the forcing of the resistent animal down the path that he didn't want, seem to me very exciting questions. They open up vistas for all sorts of further intriguing experiments. In any event, whatever we find with rats, I am going to assert that often one of the main functions of the clinician working with human beings is merely to aid the patient to discover his true drives and values so that his need-pushes and valences may come more nearly to correspond to the former. Perhaps this is the principle which Roger's non-directive therapy has most strongly emphasized and capitalized on (22).

Finally, I would like to reiterate something I said at the beginning. In the matter of the final functioning in the behavior field of need-pushes, directions and distances and valences, the neo-Gestaltists are undoubtedly right in emphasizing that the resultant behavior, or the resultant perceptual changes in the field, certainly will be in the direction of what they would call tension reduction and "better figure" in the behavior space.

Wants and cognitions as interfused activities in the concrete apprehended behavior-fields do, I believe, function according to neo-Gestalt laws (which laws have, of course, for the most part not yet been either precisely discovered or stated). But I still believe that a second system of concepts—those of drives, or needs, of values and of interconnecting beliefs which the organism carries about with him in his *personality structure* (*i.e.*, in Lewin's inner life space) and which combine with the laws of perception of external stimuli to produce the actual resultant behavior field with its inlying qualities, directions, distances, need-pushes, and valences—is necessary. I also believe that these drives, values, cathexes, and beliefs may obey more associational and less Gestalt-like laws than people like Krech and Crutchfield (13) and McLeod (15) would like to admit. I am not too insistent on this latter point, although certainly my tendency, and that of almost everybody else, to resort to such analogies as "channels" suggests associationism rather than Gestaltism. But the point I *am* insistent upon is that this problem of the drives and values and beliefs in the given personality structure is equally as necessary and important as is that of the dynamics of the immediate functioning of need-push and valence in the immediately apprehended behavior field. I shall, in short, continue to search for basic drives and continue to seek to analyze relatively enduring belief and value structures as something different from, and obeying different laws from, immediate fields and their dynamics—whether these fields be conceived phenomenologically or behavioristically. I shall persist to the end in being a drive-belief-value psychologist as well as a

behavior-field psychologist and never will regress, or advance, to being a pure all-embracing field, or neo-Gestaltish, psychologist.

REFERENCES

1. CRUTCHFIELD, R. S. Psychological distance as a function of psychological need. *J. comp. Psychol.*, 1939, **28**, 447–469.
2. FRENKEL-BRUNSWIK, E. A study of prejudice in children. *Human Relations*, 1948, **1**, 295–30.
3. ——, LEVINSON, D. J., & SANFORD, R. N. The antidemocratic personality. In *Readings in social psychology* (T. M. Newcomb and E. L. Hartley, Eds.). New York: Holt, 1947.
4. FREUD, S. *An outline of psychoanalysis.* New York: W. W. Norton & Company, 1949.
5. GOLDSTEIN, K. *The organism.* New York: American Book Company, 1939.
6. HALL, C. S. The inheritance of emotionality. *Sigma Xi Quart.*, 1935, **26**, 17–27.
7. ——, & KLEIN, S. J. Individual differences in aggressiveness in rats. *J. comp. Psychol.*, 1942, **33**, 371–383.
8. HARLOW, H. F. The formation of learning sets. PSYCHOL. REV., 1949, **56**, 51–65.
9. HULL, C. L. *Principles of behavior.* New York: D. Appleton-Century Company, 1943.
10. KENDLER, H. H., & MENCHER, H. C. The ability of rats to learn the location of food when motivated by thirst—an experimental reply to Leeper. *J. exp. Psychol.*, 1948, **38**, 82–88.
11. KOFFKA, K. *Principles of Gestalt psychology.* New York: Harcourt, Brace & Company, 1935.
12. KRECH, D. (See Krechevsky, I.) An experimental investigation of the principle of proximity in the visual perception of the rat. *J. exp. Psychol.*, 1938, **22**, 497–523.
13. ——, & CRUTCHFIELD, R. G. *Theory and problems of social psychology.* New York: McGraw-Hill Book Company, 1948.
14. LEWIN, K. The conceptual representation and the measurement of psychological forces. *Contrib. to Psychol. Theory*, 1938, **1**, No. 4.
15. MACLEOD, R. B. The phenomenological approach to social psychology. PSYCHOL. REV., 1947, **54**, 193–210.
16. MCDOUGALL, W. *An introduction to social psychology* (15th ed.). Boston: John W. Luce & Co., 1923.
17. MASLOW, A. H. The role of dominance in the social and sexual behavior of infra-human primates: IV. The determination of hierarchy in pairs and in a group. *J. genet. Psychol.*, 1936, **49**, 161–198.
18. ——. "Higher" and "lower" needs. *J. Psychol.*, 1948, **25**, 433–436.
19. MURRAY, H. A. *Explorations in personality.* New York: Oxford University Press, 1938.
20. MURPHY, G. *Personality.* New York: Harper & Brothers, 1947.
21. NOWLIS, V. Companionship preference and dominance in the social interaction of young chimpanzees. *Comp. Psychol. Monogr.*, 1941, **18**, Serial No. 85.
22. ROGERS, C. R. *Counseling and psychotherapy.* New York: Houghton Mifflin Company, 1942.
23. SANFORD, R. N., ADKINS, M. M., MILLER, R. B., COBB, E. A. *et al.* Physique, personality and scholarship, a cooperative study of school children. *Monogr. for Research in Child Develop.*, 1943, **8**, Serial No. 34.
24. SANFORD, R. N. The effects of abstinence from food upon imaginal processes: a preliminary experiment. *J. Psychol.*, 1936, **2**, 129–136.
25. SPENCE, K. W., & LIPPITT, R. An experimental test of the sign-Gestalt theory of trial and error learning. *J. exp. Psychol.*, 1946, **36**, 491–502.
26. TOLMAN, E. C. *Purposive behavior in animals and men.* New York: The Century Co., 1932.
27. ——. There is more than one kind of learning. PSYCHOL. REV., 1949, **56**, 144–155.
28. WALKER, E. L. Drive specificity and learning. *J. exp. Psychol.*, 1948, **38**, 39–49.
29. WIESNER, B. P., & SHEARD, N. M. *Maternal behaviour in the rat.* Edinburgh: Oliver and Boyd, 1933.

[MS. received June 8, 1949]

PSYCHOLOGICAL REVIEW

YEAR	VOLUME	AVAILABLE NUMBERS						PRICE PER NUMBER	PRICE PER VOLUME
1894	1	-	2	-	4	5	6	$1.00	$4.00
1895	2	-	-	3	4	5	6	$1.00	$4.00
1896	3	-	-	-	-	-	-	——	——
1897	4	1	-	-	-	-	6	$1.00	$2.00
1898	5	-	2	3	4	5	-	$1.00	$4.00
1899	6	-	-	-	-	-	6	$1.00	$1.00
1900	7	1	-	-	-	-	-	$1.00	$1.00
1901	8	1	2	-	-	-	-	$1.00	$2.00
1902	9	-	2	-	-	-	-	$1.00	$1.00
1903	10	1	2	3	-	-	-	$1.00	$3.00
1904	11	1	-	-	4	5	6	$1.00	$4.00
1905	12	1	2	3	4	5	-	$1.00	$5.00
1906	13	-	-	3	4	5	6	$1.00	$4.00
1907	14	1	2	3	-	5	-	$1.00	$4.00
1908	15	-	-	-	-	-	-	——	——
1909	16	1	-	3	4	5	6	$1.00	$5.00
1910	17	1	2	3	-	-	6	$1.00	$4.00
1911	18	1	2	3	4	5	6	$1.00	$5.50
1912	19	1	2	3	4	5	6	$1.00	$5.50
1913	20	1	2	3	4	5	6	$1.00	$5.50
1914	21	-	2	3	4	5	6	$1.00	$5.00
1915	22	1	2	3	4	5	6	$1.00	$5.50
1916	23	1	-	-	4	5	6	$1.00	$4.00
1917	24	-	-	-	4	-	-	$1.00	$1.00
1918	25	-	2	3	4	5	6	$1.00	$5.00
1919	26	1	2	3	4	5	6	$1.00	$5.50
1920	27	1	2	3	4	5	6	$1.00	$5.50
1921	28	-	2	3	4	-	6	$1.00	$4.00
1922	29	1	-	-	4	-	-	$1.00	$2.00
1923	30	1	2	3	4	5	6	$1.00	$5.50
1924	31	1	2	3	4	5	6	$1.00	$5.50
1925	32	-	2	3	-	5	-	$1.00	$3.00
1926	33	1	2	3	4	5	6	$1.00	$5.50
1927	34	1	2	3	4	5	6	$1.00	$5.50
1928	35	1	2	3	4	5	6	$1.00	$5.50
1929	36	1	2	3	4	5	6	$1.00	$5.50
1930	37	1	2	3	4	5	6	$1.00	$5.50
1931	38	1	2	3	4	5	6	$1.00	$5.50
1932	39	1	2	3	4	5	6	$1.00	$5.50
1933	40	1	2	3	4	5	6	$1.00	$5.50
1934	41	1	2	3	4	5	6	$1.00	$5.50
1935	42	1	2	3	4	5	6	$1.00	$5.50
1936	43	1	2	3	4	5	6	$1.00	$5.50
1937	44	1	2	3	4	5	6	$1.00	$5.50
1938	45	1	2	3	4	5	6	$1.00	$5.50
1939	46	1	2	3	4	5	6	$1.00	$5.50
1940	47	1	2	3	4	5	6	$1.00	$5.50
1941	48	1	2	3	4	5	6	$1.00	$5.50
1942	49	1	2	3	4	5	6	$1.00	$5.50
1943	50	1	2	3	4	5	6	$1.00	$5.50
1944	51	-	2	3	4	5	6	$1.00	$5.00
1945	52	1	2	3	4	5	6	$1.00	$5.50
1946	53	1	-	3	4	5	6	$1.00	$5.00
1947	54	1	2	3	4	5	6	$1.00	$5.50
1948	55	1	2	3	4	5	6	$1.00	$5.50
1949	56	y Subscription, $5.50						$1.00	——
								List price, Volumes 1 through 55	$241.50
								30% Discount	72.45
								Net price, Volumes 1 through 55	$169.05

Information about the Psychological Review: from 1894 to 1908, many numbers are out of print, as shown in the table. After 1908, twenty-three numbers are out of print. After 1925, two numbers are out of print.

The journal has been published with six numbers a year throughout its history.

Information about prices: the Psychological Review has the uniform price of $5.50 per volume and $1.00 per issue. For incomplete volumes, the price is $1.00 for each available number. For foreign postage, $.25 per volume should be added. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of $ 50.00 and over
20% on orders of $100.00 and over
30% on orders of $150.00 and over

Current subscriptions and orders for back numbers should be addressed to

AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1515 Massachusetts Avenue, N. W. Washington 5, D. C.

PSYCHOLOGY *Ready January, 1950*

Henry E. Garrett, Head of the Department of Psychology, Columbia University

Readable

This text is written in a highly interesting style which is brightened by occasional touches of humor. It is less "technical" in the old-fashioned sense than the usual type of psychology text and yet it covers all the essential information required in a beginning course in psychology. An overly detailed treatment of physiology is avoided. Employing a practical approach, this up-to-date book helps students to meet problems of everyday living, such as studying effectively, getting a job, and understanding himself and others.

Teachable

The logical organization and clarity of style appeal to teachers and students alike. The text matter is illustrated by approximately sixty-five charts and graphs, which in many instances take the place of long verbal discussions. At the end of each chapter there are experiments for the student to make, topics for him to think about and discuss, and outside reading suggestions which refer to specific sections of the books.

Approx. 336 pages

American Book Company

JOURNAL OF CONSULTING PSYCHOLOGY

Issued bi-monthly *$5.00 per year*

American Psychological Association

1515 Massachusetts Avenue, N.W.

Washington 5, D.C.

PSYCHOLOGICAL REVIEW

VOL. 57
NOS 1-6

1950

PUB. 279

UNIVERSITY MICROFILMS
ANN ARBOR, MICHIGAN. 1951